Analysis of Convex Functions

Shivam Raj

A Thesis submitted to

Indian Institute of Technology Hyderabad

In Partial Fulfillment of the Requirement for

The degree of Master of Science

In the guidance of

Dr. Tanmoy Paul

Department of Mathematics

May 2018

Declaration

I declare that this written submission represents my project work, and where ideas or words of others have been included, I have adequately referenced the original sources. I own the mistake, if any, crept into this report and do not hold anybody or any reference responsible for such mistakes.

Shivam Raj

(Signature)

SHIVAM RAJ

(Shivam Raj)

MA16MSCST11015

(Roll No.)

Approval Sheet

This Thesis entitled **Analysis of Convex Functions** by Shivam Raj is approved for the degree of Master of Science from IIT Hyderabad

(Dr. Tanmoy Paul)Adviser
Dept. of Mathematics
IITH

11/5/18

(Dr. C.S Sastry)Faculty-Adviser
Dept. of Mathematics
IITH

9/5/18

(Dr. J. Balasubramaniam)H.O.D
Dept. of Mathematics
IITH

Abstract

Convexity is an old subject in mathematics. The first specific definition of convexity was given by Herman Minkowski in 1896. Convex functions were introduced by Jensen in 1905. The concept appeared intermittently through the centuries, but the subject was not really formalized until the seminal 1934 tract *Theorie der konverxen Korper* of Bonneson and Fenchel.

Today convex geometry is a mathematical subject in its own right. Classically oriented treatments, like the work done by Frederick Valentine form the elementary definition, which is that a domain K in the plane or in $\mathbb{R}^N$ is convex if for all $P, Q \in K$, then the segment $\overline{PQ}$ connecting P to Q also lies in K. In fact this very simple idea gives forth a very rich theory. But it is not a theory that interacts naturally with mathematical analysis. For analysis, one would like a way to think about convexity that is expressed in the language of functions and perhaps its derivatives.

Our goal in this thesis is to present and to study *convexity* in a more analytic way. Through Chapter 1, Chapter 2 and Chapter 3, I have tried to point out the important role of convex sets and its associated convex functions in Mathematical Analysis. Chapter 1 is devoted to Convex sets and some geometric properties achieved by these objects in finite Euclidean spaces. The emphasis is given on establishing a criteria for convexity. Various useful examples are given, and it is shown how further examples can be generated from these by means of operations such as addition or taking convex hulls. The fundamental idea to be understood is that the convex functions on $\mathbb{R}^n$ can be identified with certain convex subsets of R^{n+1} (their epigraphs), while the convex sets in $\mathbb{R}^n$ can be identified with certain convex functions on $\mathbb{R}^n$ (their indicators). These identifications make it easy to pass back and forth between a geometric approach and an analytic approach. Chapter 2 begins with idea of convexity of functions in a finite dimensional space. Convex functions are an important device for the study of extremal problems. They are also important analytic tools. The fact that a convex function can have at most one minimum and no maxima is a notable piece of information that proves to be quite useful. A convex function is also characterized by the non negativity of its second derivative. This useful information interacts nicely with the ideas of calculus. We relate convex functions to an elegant characterisation of Gamma functions by Bohr Mollerup Theorem. Chapter 3 provides an introduction to convex analysis, the properties of sets and functions in infinite dimensional space. We start by taking the convexity of the epigraph to be the definition of a convex function, and allow convex functions to be extended -real valued. One of the

main themes of this chapter is the maximization of linear functions over non empty convex sets. Here we relate the subdifferential to the directional derivative of a function. There are several modern works on convexity that arise from studies of functional analysis. One of the nice features of the analytic way of looking at convexity is the Bishop-Phelps Theorem, it says that in a Banach Space, a convex function has a subgradient on a dense subset of its effective.

Acknowledgements

First and foremost I want to thank my adviser **Dr. Tanmoy Paul**. It has been an honor to be his student. He has taught me consciously. I appreciate his contributions of time and ideas, to make my M.Sc. project experience productive. The joy and enthusiasm he has for his research was contagious and motivational for me. I am thankful for the excellent examples he has provided as a successful professor.

I am also grateful to Soumitra Daptari for the help he provided me during this project. Lastly, I would like to thank my family for all their love and encouragement,they are the ones my who instilled in me my love for mathematics and they have always supported me in all my pursuits. Thank you.

Shivam Raj,
IIT Hyderabad.

May 2018.

Contents

Dedicated to my parents

Chapter 0

Introduction and Definitions

0.1 To begin with

Convexity is a simple and natural notion which can be traced back to Archimedes (250 B.C.) in the connection with his famous estimate of the value of π (by using inscribed and circumscribed regular polygons). He noticed the important fact that the perimeter of a convex figure is smaller than the perimeter of any other convex figure surrounding it.

As a matter of fact, we experience convexity all the time and in many ways. The most prosaic example is our upright position, which is secured as long as the vertical projection of our center of gravity lies inside the convex envelope of our feet. Also, convexity has a great impact on our everyday life through numerous applications in industry, business, medicine and art. So do the problems of optimum allocation of resources and equilibrium of non-cooperative games.

The recognition of the subject of convex functions as one that deserves to be studied in its own right is generally ascribed to J. L. W. V. Jensen. However, he was not the first to deal with such functions. Ch. Hermite, O. Holder and O. Stolz were among the other stalwarts. During the twentieth century, there was intense research activity and significant results were obtained in *Geometric Functional Analysis, Mathematical Economics, Convex Analysis, Nonlinear Optimization*. A classic book by G. H. Hardy, J. E. Littlewood and G. Polya played a large role in the realm of the subject of convex functions.

Roughly speaking, there are two basic properties of convex functions that make them so widely used in theoretical and applied mathematics.

- The Maximum is attained at a boundary point.
- Any local minimum is a global one. Moreover, a strictly convex function admits at most one minimum.

The modern viewpoint of convex functions entails a powerful and elegant interaction between analysis and geometry. *Geometric Functional Analysis* is the branch where people studies the extent to which the properties possessed by finite dimensional spaces generalize to infinite dimensional spaces. In finite dimensional vector space there is only one natural linear topology. In that topology every linear functional is continuous, any convex functional is also continuous (at least on the interior of the domain). Convex hull of a compact set is compact and any two disjoint closed convex set can be separated by hyperplanes. On the other hand in an infinite dimensional space there are more than one topology which *respect* the linear structures of the space, the topological dual (ie. the set of all continuous linear functionals) depends on the topology endowed on this space. In infinite dimensional space the convex functionals not necessarily continuous, convex hull of a compact set not necessarily compact and disjoint closed convex sets cannot generally separated by hyperplanes. However, with the right topology and perhaps some additional assumptions, each of these results has an appropriate infinite dimensional version.

Our aim in this project is to relate the analytical structure of convex functions with the structure of underlying space by means of the geometry of the space. We first extract some geometric properties of convex sets in a linear space and then we switch our attention to the convex functions.

0.2 A few necessary ingredients

$\mathbb{N}$ represents set of natural numbers.

$\mathbb{Z}$ represents set of integers.

$\mathbb{R}$ represents set of real numbers.

$\hat{f}(x)$ define $\inf_{x\in K}\{g(x) : g \geq f \quad \text{and} \quad g \quad \text{is} \quad \text{affine}, K(\subseteq \mathbb{R}^n)\}$.

$\check{f}(x)$ define $\sup_{x\in K}\{g(x) : g \leq f \quad \text{and} \quad g \quad \text{is} \quad \text{affine}, K(\subseteq \mathbb{R}^n)\}$.

$[f = \alpha]$ denote the level set $\{x : f(x) = \alpha\}$

$[f < \alpha]$ denote the level set $\{x : f(x) < \alpha\}$

Definition 0.2.1 (Convex Set)**.** Let K be a subset of a vector space X Then K is said to be convex if for all $x, y \in K$ and $\lambda \in [0, 1]$ it satisfies the following condition:

$$\{(1-\lambda)x + \lambda y : \lambda \in [0,1]\} \in K$$

.

Let K be a convex subset of a vector space X and f is a real valued function on K.

Definition 0.2.2 (Convex Function)**.** The function f is said to be convex on K if for all $x_1, x_2 \in K$ and $\lambda \in [0, 1]$ it satisfies the inequality:

$$f(x) \leq \lambda f(x) + (1-\lambda)f(x). \tag{1}$$

If strict inequality holds whenever $x_1 \neq x_2$ and $\lambda \in (0, 1)$, call f strictly convex on K.

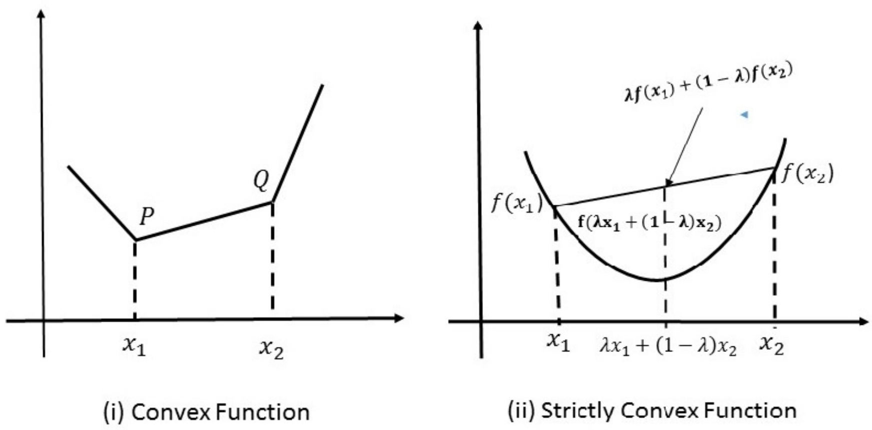


FIGURE 1: Convex Function

Definition 0.2.3 (Concave Function)**.** A function f is called a convex function if for all $x_1, x_2 \in K$ and $\lambda \in [0, 1]$ then

$$f(x) \geq \lambda f(x_1) + (1-\lambda)f(x_2) \tag{2}$$

Definition 0.2.4 (Epigraph of a function)**.** The $epi(f)$ of a function $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ is the set of points laying on or above its graph.

$$epi(f) = \{(x, y) : x \in dom(f) i.e, x \in K, y \geqslant f(x)\} \subset \mathbb{R}^{n+1}$$

Remark 0.2.5. (i) It is also known as **Supergraph** of a function f.

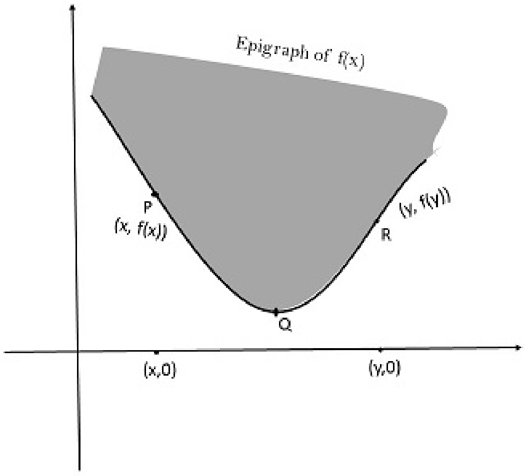


FIGURE 2: Epigraph of $f(x)$

(ii) **Strictly epigraph** is the epigraph with the graph of the function removed from the Supergraph.

$$epi_s(f) = \{(x, y) : x \in \mathbb{R}^n, K \in \mathbb{R}, y > f(x)\} \subseteq \mathbb{R}^{n+1}$$

Definition 0.2.6 (Convex hull)**.** If $S \subset \mathbb{R}^n$, the convex hull of S is defined to be the intersection of all convex sets containing S.

$$Conv(S) = \{\sum_{i=1}^{n} t_i u_i : n \in \mathbb{N}, u_1, u_2, ..., u_n \in S, t_i \in [0, 1], 1 \le i \le n \quad \text{and} \quad \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{t_i} = 1\}.$$

In other words, $Conv(S)$ consists of all possible convex combination of points in S.

Definition 0.2.7 (Common Transversal). Let S be a family of sets in $\mathbb{R}^2$. A line which meets every member of S is called a common transversal of S.

Definition 0.2.8 (Hyperplane). Let X is a vector space. A hyperplane is a set of the form $[f = \alpha]$, where f is a non-zero linear functional on X.

- A hyperplane defines two **strict half spaces** $[f < \alpha]$ and $[f > \alpha]$ and two **weak half spaces** $[f \le \alpha]$ and $[f \ge \alpha]$.

Definition 0.2.9 (Polyhedron). A set in a vector space is Polyhedron if it it the intersection of finitely many weak half spaces.

Definition 0.2.10 (Extreme Point). Let K be a convex set. A point $x \in K$ is an extreme point of K if $x = tx_1 + (1 - t)x_2, (0 < t < 1, x_1, x_2 \in K)$ then $x = x_1 = x_2$. In other words, x cannot be an interior point of a line segment in K.

Definition 0.2.11 (Face). A convex subset F of K is said to be face if for any $x, y \in K$, $\lambda x + (1 - \lambda)y \in F$ for some $\lambda \in (0, 1)$ implies $x, y \in F$.

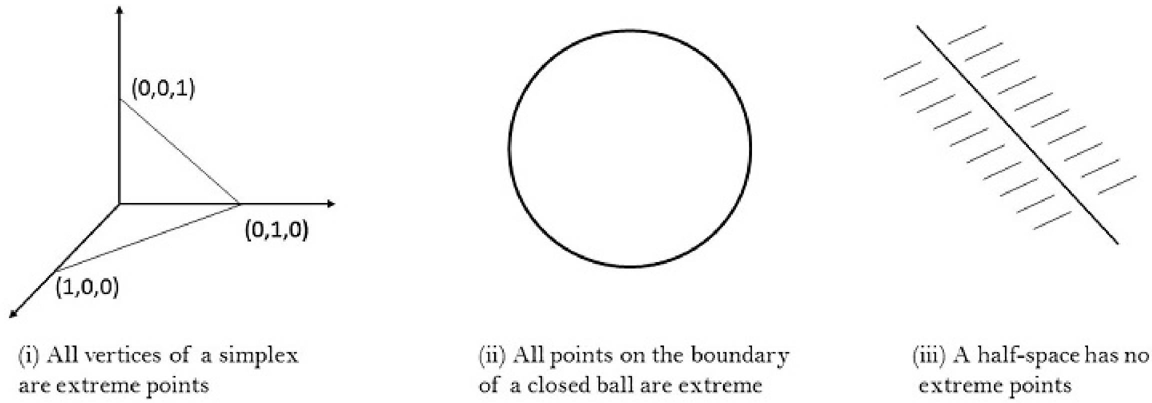


FIGURE 3: Examples of extreme points

Definition 0.2.12. (**Affine Function) A function $f : X \to \mathbb{R}$ on a vector space is affine if it is of the form**

$$f(x) = x^*(x) + c$$

for some linear function $x^* \in X^*$ an some real c.

- Every linear functional is affine and every affine function is both convex and concave.

Definition 0.2.13 (**Topological space).** Let X be a non-empty set. A family τ of subsets of X is called a topology on X if it satisfies following two conditions:

(i) $\phi, X \in \tau$

(ii) the union of every family of sets in τ is a set in τ.

(iii) the intersection of every finite family of sets in τ is a set in τ.

If τ is an topology on X, then the pair (X, τ) is called **topological space**.

Definition 0.2.14 (**Hausdorff Space).** A hausdorff space is a topological space in which each pair of distinct points can be separated by open sets. In the sense that they have disjoint neighborhoods.

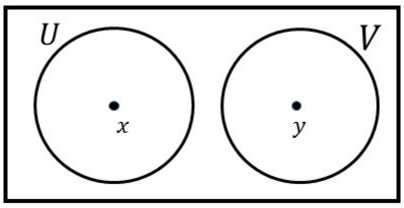


FIGURE 4: The points x and y, separated by their respective neighborhoods U and V

- The real numbers are Hausdorff space.

Definition 0.2.15 (Topological vector space). A topology τ on a vector space X is called **linear topology** if the operations addition and scalar multiplication are τ-continuous, That is, If $(x, y) \mapsto x+y$ from $X \times X \to X$ and $(\alpha, x) \mapsto \alpha x$ from $\mathbb{R} \times X \to X$ are continuous. Then (x, τ) is called topological vector space (or, Linear topological space).

- A topological vector space need not be a Hausdorff space.
- All Banach Spaces and Hilbert Spaces are topological vector space.

Definition 0.2.16 (Locally Convex Space). A topological vector space is locally convex, or is a locally convex space. If every neighborhood of zero includes a convex neighborhood of zero.

Definition 0.2.17 (Positive Homogeneity). A function $f : V \backslash \{0\} \to \mathbb{R}$ is positively homogeneous of degree k if

$$f(\alpha x) = \alpha^k f(x) \text{ for all } \alpha > 0$$

Here k can be any real number.

Chapter 1

Convex Sets

1.1 Geometry of Convex Sets

Recall 1.1.1 (Convex Set)**.** Let set $K \subset \mathbb{R}^n$. Then K is convex if the line segment joining any two points of K is contained in K.

FIGURE 1.1: Convex and Concave set

Example 1.1.2. Following are few examples of convex set:

(i) $B(x_0, r) := \{x \in \mathbb{R}^n : \|x - x_0\| < r\}$, the open ball of radius r with center x_0, is convex as is the closed ball $\bar{B}(x_0, r) := \{x \in \mathbb{R}^n : \|x - x_0\| \leq r\}$.

(ii) The open and closed half-space $\{x : x.a < c\}, \{x : x.a \leq c\}, \{x : x.a > c\}, \{x : x.a \geq c\}$ are all convex.

(iii) The intersection of any collection of convex sets is convex.

(iv) A set $H = \{x \in \mathbb{R}^n : \alpha^T \geq \beta\}$, where $\alpha \in \mathbb{R}^n, \alpha \neq 0, \beta \in \mathbb{R}$ is convex set.

Explanation:

(i) Since $\|\lambda x + (1-\lambda)y - x_0\| = \|\lambda(x - x_0) + (1-\lambda)(y - x_0)\| \leq \lambda\|x - x_0\| + (1-\lambda)\|y - x_0\| < \lambda r + (1-\lambda)r = r$. Hence $\lambda x + (1-\lambda)y \in B(x_0, r)$ for all $x, y \in B(x_0, r)$. So,$B(x_0, r)$ is convex. Similarly, we can show that $\bar{B}(x_0, r)$ is convex.

(ii) Let $H_1 = \{x : x.a < c\}$. Take to points $x, y \in H_1$. we can easily see that $(\lambda x + (1-\lambda)y).a < c$. i.e, $(\lambda x + (1-\lambda)y) \in H_1$. So, H_1 is convex. Similarly, we can show for other three half-spaces.

(iii) Consider $X = \bigcap_{i \in I} k_i$, where k_i are convex set. Let $x, y \in X$. Since $k_1, k_2, ...$ are convex sets. So, $\lambda x + (1-\lambda)y$ belong to K_i for each fix i then $\lambda x + (1-\lambda)y$ also belongs to X. Hence X is convex.

(iv) Take $x, y \in H$.Consider $z = (1-\lambda)x + \lambda y$ for any $\lambda \in (0,1)$.Then $\alpha^T = \alpha^T((1-\lambda x + \lambda y) = (1-\lambda)\alpha^T x + \lambda \alpha^T y) \geq (1-\lambda)\beta + \lambda\beta = \beta$ Thus, $\alpha^T \geq \beta$.Hence $z \in H$ and hence H is convex set.

Some typical results in Convexity theory in $\mathbb{R}^n$:

Theorem 1.1.3 (Helly's First Theorem)**.** Let $K_i, i = 1, 2, ..., N, (N \geq n+1)$, be convex sets in $\mathbb{R}^n$. If any $(n+1)$ of them have a non-void intersection, then $\bigcap_{i=1}^{N} K_i \neq \emptyset$.

Proof. By induction on N. For $N = n+1$, the assertion is correct. Suppose that the statement of the theorem holds for $N-1 \geq n+1 (i.e, N \geq n+2)$ convex sets. Then by the induction hypothesis, there exist points

$$X_j \in \bigcap_{j=1, j\neq i}^{N} K_j, i = 1, 2, ..., N.$$

Take $N-1(\geq n+1)$ vectors $(X_1 - X_N, X_2 - X_N, ..., X_N - 1 - X_N)$ being linearly dependent, there exist scalars $C_1, C_2, ..., C_{N-1}$ not all zero such that

$$\sum_{i=1}^{N-1} (X_i - X_N) = 0$$

. i.e, scalars $\{\lambda_i\}_{i=1}^{N}$ not all zero such that

$$\sum_{i=1}^{N} \lambda_i X_i = 0 = \sum_{i=1}^{N} \lambda_i.$$

(put $\lambda_i = C_i, 1 \leq i \leq N-1$ and $\lambda_N = -(C_1 + C_2 + ... + C_N - 1))$.
without loss of generality, we may assume that $\lambda_1 \geq 0, \lambda_2 \geq 0, ..., \lambda_N \geq 0$ and $\lambda_{r+1} \leq 0, ..., \lambda_N \leq 0$. Then $\sum_{i=1}^{r} \lambda_i = -\sum_{i=r+1}^{N} \lambda_i$ and we see that the point

$$Y := \frac{\sum_{i=1}^{r} \lambda_i X_i}{\sum_{i=1}^{r} \lambda_i}$$

belongs to $\bigcap_{j=r+1}^{N} K_j$ as every $X_i, i = 1, 2, ..., r$ belongs to every $K_j, j = r+1, ..., N$. But we have also that

$$Y = \frac{\sum_{i=r+1}^{N}(-\lambda_i)X_i}{\sum_{i=r+1}^{N}(-\lambda_i)}$$

which shows that $Y \in \bigcap_{j+1}^{N} K_j$.Hence $Y \in \bigcap_{j=1}^{N} K_j \neq \emptyset$. □

Theorem 1.1.4 (Second Helly's Theorem)**.** Let $\mathscr{K} = \{K_i : i \in I\}$ be a family (positively infinite) of compact convex sets in $\mathbb{R}^n$. If any (n+1) members of $\mathscr{K}$ have non-empty intersection, then the intersection $\bigcap_{i \in I} K_i \neq \emptyset$.

Let us consider one simple application of Helly's Theorem :

Theorem 1.1.5 (Transversal theorem of Santalo)**.** Let $\mathscr{S} = \{S_1, S_2, ..., S_m\}$ be a family of parallel line segments $S_i \subseteq \mathbb{R}^2$. Some line segments may consist of single points. If every there members of $\mathscr{S}$ have a common transversal.

Proof. Without loss of generality that the line segment are vertical $\mathscr{S}$. i.e, of the form

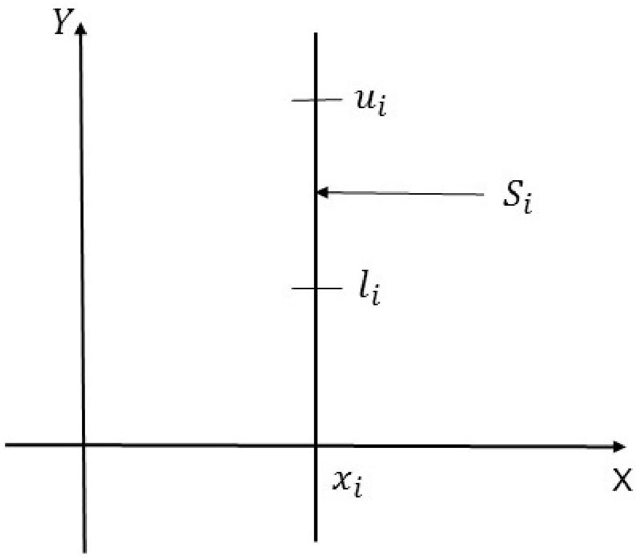


FIGURE 1.2

$$S_i = \{(x, y) := x_i, \quad l_i \leq y \leq u_i\}.$$

with each line segment s_i associates the convex set

$$C_i = \{(a, b) : l_i \leq ax_i + b \leq u_i\}.$$

By induction, every these sets C_i have a non-empty intersection and we can apply Helly's first theorem to get the required conclusion. □

A basic result in this theory is

Theorem 1.1.6 (Caratheodry's Theorem)**.** If S is a compact convex set in $\mathbb{R}^n$. Then each $x \in S$ can be expressed as a convex combination of at most $(n+1)$ extreme points of S.

To prove Caratheodry's result, we first prove a lemma:

Lemma 1.1.7. Let $x \in conv(S')$, where $S' \subset \mathbb{R}^n$ then x is a convex combination of at most $(n+1)$ point of S'.

Proof. $x = \sum_{j=1}^{m} t_j x_j$ be a convex combination of $x_1, x_2, ..., x_m \in S'$, where $m > n+1$ (Assume all $t_j > 0$). We show that x is a convex combination of $(m-1)$ points of S'. Since,$m-1 > n$,there exist $c_1, c_2, ..., c_m$ not all 0 such that

$$\sum_{j=1}^{m} c_j = 0, \quad \sum_{j=1}^{m} c_j x_j = 0$$

Let $s_j = t_j - -\alpha c_j$ for j=1,2,3,.....,m
where $\dfrac{1}{\alpha} = \max\{\dfrac{c_1}{t_1}, ..., \dfrac{c_m}{t_m}\}$ and consequently,

$$x = \sum_{j \neq K} s_j x_j, \quad \sum_{j \neq K} s_j = 1$$

where $s_k = 0$ for some k, i.e x is a convex combination of the $m-1$ points $x_1, ..., x_{k-1}, x_{k+1}, ..., x_m$. Either $m-1 = n+1$ or else we continue the same argument to show that x is a convex combination of $m-2$ points of S'. Continuing, we find that x can be expressed as a convex combination of $n+1$ or fewer points of S'. □

Before we proceed to the proof of the above result, there are a couple of facts – separation of convex sets that are summarized in the following well-known theorem.

Theorem 1.1.8. Suppose A, B are disjoint non-empty convex sets in $\mathbb{R}^n$.

(i) If A is open, there exist $a \in \mathbb{R}^n$ and $r \in \mathbb{R}$ such that $a.x < r \leq a.y$ for every $x \in A$ and for every $y \in B$. In Figure 1.3(i),

(ii) If A is compact and B is closed then there exist $a \in \mathbb{R}^n, r_1, r_2 \in \mathbb{R}$ such that $a.x < r_1 < r_2 < a.y$ for every $y \in B$. (This is called strict separation of A and B). In Figure 1.3(ii),

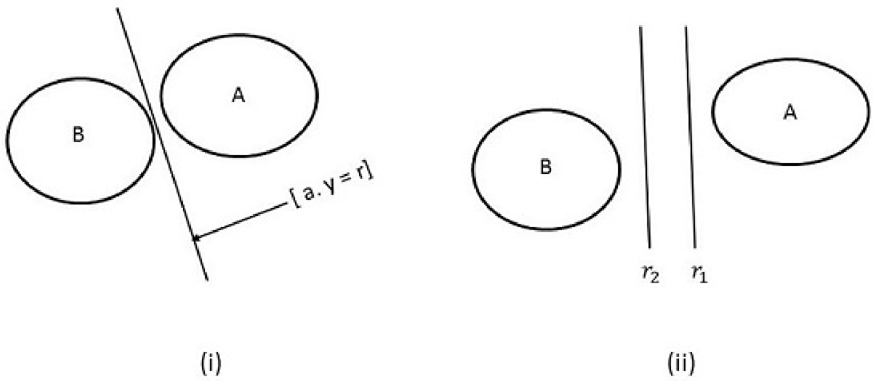


FIGURE 1.3

Recall 1.1.9. A **hyperplane** H in $\mathbb{R}^n$ is a set of the form $x : a.x = \alpha$ where $\alpha \in \mathbb{R}$. With the aid of the above results, one can easily prove another standard fact:
If y is a boundary point of a closed convex set K, then K has a **supporting hyperplane** P which contains y. This means $P \cap K \neq \emptyset$, $\quad P \cap K$ is convex and contains only boundary points of K.

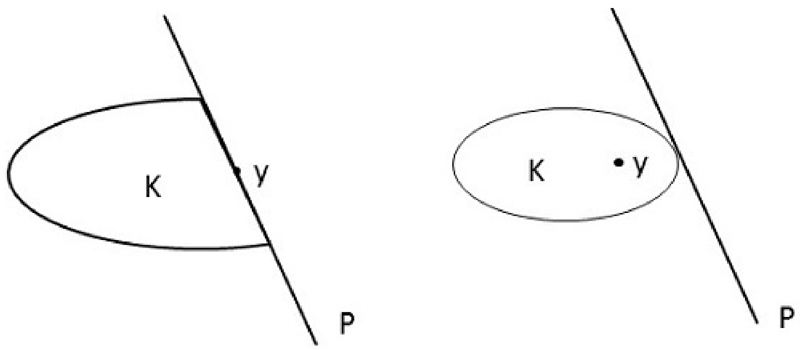


FIGURE 1.4

Completion of proof : Use induction on n. For $n = 1$, $\quad S = [a, b]$ and the result is obvious. Assume the theorem holds for $n - 1$. Let $S \subset \mathbb{R}^n$ with $dimS = n$ (this is equivalent to saying $int(s) = \emptyset$). If $x \in \delta S$ (=boundary of S), then by the separation theorem, there is a supporting hyperplane H of s such that $x \in H \cap S$. By the induction hypothesis, x is convex combination of at most n extreme points of $H \cap S$ and these extreme points must also be extreme in S.(verify)

If $z \in int(S)$, choose z_0 extreme in S and draw the line segment through z_0 and z which hits S in the boundary point y. See Figure (1.5),

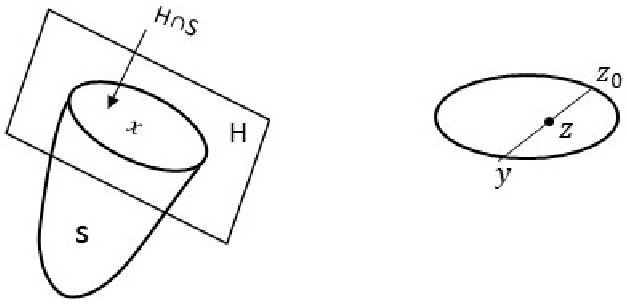


FIGURE 1.5

Then by the first paragraph, $y = \sum_{i=1}^{n} t_i z_i$ as a convex combination of extreme points. As $z = t_0 y + (1 - t_0) z_0$ for $t_0 \in (0, 1)$, we must have

$$z = \sum_{i=1}^{n} t_0 t_i z_i + (1 - t_0) z_0$$

which is a convex combination of $(n + 1)$ extreme points $z_1, z_2, ..., z_n$ and this completes the proof.

Remark 1.1.10. It is clear that we have not use the lemma (1.1.7) to proof Caratheodory's theorem. However the lemma is of independent interest and is also attributed to Caratheodory.

Chapter 2

Convex Functions

2.1 A few basic results of Convex function in $\mathbb{R}$

In this section we are going to discuss basics results of convex function in $\mathbb{R}$ and assume that $\mathbb{E}$ be a convex subset of $\mathbb{R}$.

Definition 2.1.1 (Convex function on $\mathbb{R}$)**.** A function$f : \mathbb{E} \to \mathbb{R}$ is said to be convex if for all $\lambda \in [0, 1]$ and for all $x, y \in \mathbb{E}$, the following inequality holds:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)y \tag{2.1}$$

Example 2.1.2. Following are few examples of convex function on $\mathbb{R}$

(i) A function $f : \mathbb{R} \to \mathbb{R} \geq 0$ and given by $f(x) = x^2$ Then $f(x)$ is a convex function.

(ii) If $f : \mathbb{R} \to \mathbb{R}$ and given by $f(x) = |x|^p, \quad 1 \leq p < \infty$ then f is a convex function.

Explanation:

(i) If $f(x)$ is convex then it satisfy the condition of convexity;

$$\begin{aligned}
\text{i.e,} \quad & (\lambda x_1 + (1 - \lambda)x_2)^2 \leq \lambda x_1^2 + (1 - \lambda)x_2^2 \\
\text{i.e,} \quad & \lambda x_1^2 + (1 - \lambda)^2 x_2^2 + 2\lambda(1 - \lambda)x_1 x_2 \leq \lambda x_1^2 + (1 - \lambda)x_2^2 \\
\text{i.e,} \quad & \lambda(1 - \lambda)x_1^2 + \lambda(1-\lambda)x_2^2 + 2\lambda(1 - \lambda)x_1 x_2 \leq 0 \\
\text{i.e,} \quad & \lambda(1 - \lambda)(x_1 + x_2)^2 \leq 0
\end{aligned}$$

Since, $\lambda \geq 0, \lambda - 1 \leq 0$ and $(x_1 + x_2)^2 \geq 0$. Thus above inequality is true. Hence $f(x) = x^2$ is convex function.

(ii) By differentiating $f(x)$ we get,

$$f'(x) = \begin{cases} p|x|^{p-1} & \text{if} x \geq 0 \\ -p|x|^{p-1} & \text{if} x < 0 \end{cases} \tag{2.2}$$

Again differentiate (2.2) we get,

$$f''(x) = \begin{cases} p(p-1)|x|^{p-2} & \text{if} x \geq 0 \\ p(p-1)|x|^{p-2} & \text{if} x < 0 \end{cases} \tag{2.3}$$

It is clear that $f''(x) \geq 0$ because $p(p-1) > 0$ and $|x|^{p-2} \geq 0$. So,by theorem (2.1.5) $f(x)$ is convex .

Geometric Meaning: (i) The line joining $(y, f(y))$ and (x, f (x)) in $\mathbb{R}^2$,for $0 \leq \lambda \leq 1$ is given by $\lambda(x, f(x)) + (1-\lambda)(y, f(y)) = (\lambda x + (1-\lambda)y, \lambda f(x) + (1-\lambda)f(y))$. Then the

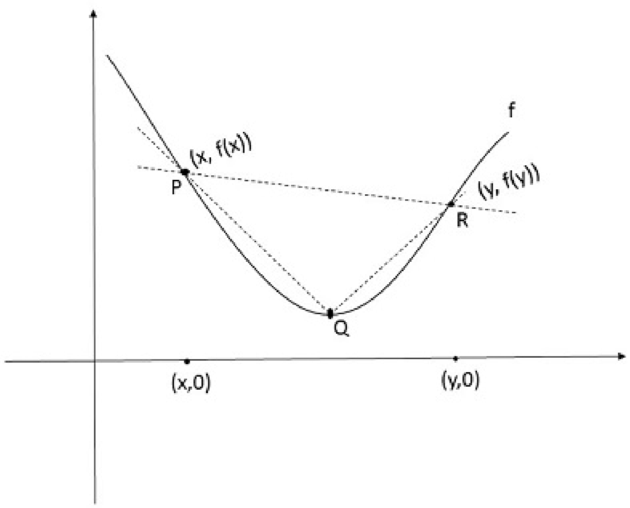


FIGURE 2.1

y-coordinates of points on this line are given by $\lambda f(x) + (1-\lambda)f(y)$. Thus the geometric meaning of the definition is that the line segment joining $(x, f(x))$ and $(y, f(y))$ is never below the graph of the function. See Fig.(2.1).
(ii) If P,Q and R be three points on the graph of f with Q between P and R, then Q is on or below chord PR.

Proposition 2.1.3. Let $f : [a, b] \to \mathbb{R}$ is convex if and only if any of the following condition is satisfied: for all $x, u, y \in [a, b]$ with $a \leq x < u < y \leq b$

(i) $\dfrac{f(u) - f(x)}{u - x} \leq \dfrac{f(y) - f(x)}{y - x}$

(ii) $\dfrac{f(u)-f(x)}{u-x} \leq \dfrac{f(y)-f(u)}{y-u}$

Proof. Given that f is convex on E and $x < u < y$. Then we can write $u = \lambda x + (1-\lambda)y$ where $\lambda = \dfrac{y-u}{y-x} \in [0,1]$.Since f is convex

$$f(u) \leq \frac{y-u}{y-x} f(x) + \frac{u-x}{y-x} f(y) \tag{2.4}$$

$$\begin{aligned}
&\Leftrightarrow f(u) - f(x) \leq \frac{x-u}{y-x} f(x) + \frac{u-x}{y-x} f(y) \\
&\Leftrightarrow f(u) - f(x) \leq \frac{u-x}{y-x}[f(y) - f(x)] \\
&\Leftrightarrow \frac{f(u)-f(x)}{u-x} \leq \frac{f(y)-f(x)}{y-x}
\end{aligned} \tag{2.5}$$

Hence the (i) inequality is proved.we have from equation(2.4),

$$f(u) \leq \lambda f(x) + (1-\lambda) f(y)$$

$$\begin{aligned}
&\Leftrightarrow \lambda[f(x) - f(y)] + [f(y) - f(u)] \geq 0 \\
&\Leftrightarrow \frac{y-u}{y-x}[f(y) - f(x)] \leq f(y) - f(u) \\
&\Leftrightarrow \frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(u)}{y-u}
\end{aligned} \tag{2.6}$$

From equation (2.5) and (2.6) we get,

$$\frac{f(u)-f(x)}{u-x} \leq \frac{f(y)-f(u)}{y-u} \tag{2.7}$$

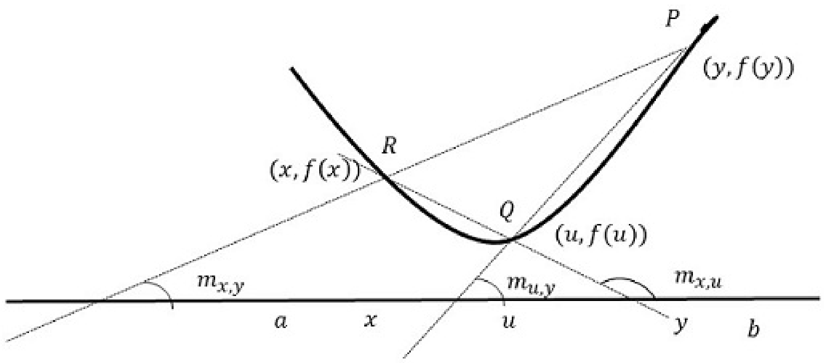


FIGURE 2.2

This complete the proof. It follows form the equation (2.5) and (2.7), we get an important and generalized result for all $x, u, y \in [a, b]$,

$$\frac{f(u) - f(x)}{u - x} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(u)}{y - u} \tag{2.8}$$

□

Note 2.1.4. Let us define $m_{u,v} := \dfrac{f(v) - f(u)}{v - u}$ for any $u, v \in [a, b]$ this gives us slope of the line joining $(u, f(u))$ and $(v, f(v))$. The geometric meaning of equation (2.8) gives us

$$\text{slope(PQ)} \leq \text{slope(PR)} \leq \text{slope(QR)}$$

see the above picture (2.2).

Theorem 2.1.5. Let $f : \mathbb{E} \to \mathbb{R}$ be twice differentiable function. Then the following are equivalent

(i) f is convex.

(ii) f' is monotonically increasing on (a, b).

(iii) f'' is positive i.e.,$f''(x) \geq 0\ \forall x \in (a, b)$.

(iv) $f'(x_0)(y - x_0) \leq f(y) - f(x_0), \quad \forall x_0, y \in \mathbb{E}$

Proof. (i)$\Rightarrow$ (ii) Assume that f is differentiable and convex on (a, b). Let $x, y \in (a, b)$. Let $x, y \in (a, b)$ with $x < y$. By the previous proposition, for any $u \in [x, y]$. we have,

$$m_{x,u} \leq m_{x,y} \leq m_{u,y}$$

Letting u decreases to x in the first inequality yields

$$f'_-(x) \leq m_{x,y}$$

. Similarly u increases to y in the second inequality yields

$$m_{x,y} \leq f'_+(x)$$

. Thus we get

$$f'_-(x) \leq f'_+(x) \tag{2.9}$$

Since $f'_-(x) = f'_+$ for all $x \in (a, b)$. Hence we get,$f'(x) \leq f'(y)$
That is, f' is increasing.

(ii)⇒(i): Assume that f' is monotonically increasing on (a, b). Let $(x, y) \in (a, b), x < y$ and $\lambda \in (0, 1)$ and $S = \lambda x + (1 - \lambda)y$.Thus, $x < s < y$.then by Mean Value theorem there exists c_1 such that $x < c_1 < s$ and

$$f'(c) = \frac{f(s) - f(s)}{s - x}$$

and there exists c_2 such that$x < c_2 < y$ and $f'(c_2) = \dfrac{f(y) - f(s)}{y - s}$.
Since, f is monotonically increasing and $c_1 < c_2$, this implies

$$\begin{aligned}
\frac{f(s) - f(x)}{s - x} &\leq \frac{f(y) - f(x)}{y - s} \\
\frac{\lambda x + (1 - \lambda)y - f(x)}{\lambda x + (1 - \lambda)y - x} &\leq \frac{f(y) - f(\lambda x + (1 - \lambda)y}{y - (\lambda x + (1 - \lambda)y} \\
\lambda f(\lambda x + (1 - \lambda)y) - \lambda f(x) &\leq (1 - \lambda)f(y) - (1 - \lambda)f(\lambda x + (1 - \lambda)y) \\
\text{or} \quad \mathrm{f(\lambda x + (1 - \lambda)y)} &\leq \lambda f(x) + (1 - \lambda)f(y)
\end{aligned}$$

So, f is convex.

(iii)⇒(ii): Let $f''(x) \geq 0$ on $\forall(a, b)$. Take any two points x_1, x_2 such that $x_1 < x_2 \quad \forall x_1, x_2 \in (a, b)$. Then by Mean Value Theorem, there exist $c \in (x_1, x_2)$ such that

$$f''(c) = \frac{f'(x_2) - f'(x_1)}{x_2 - x_1}$$

Since $f''(x) \geq 0 \quad \forall x \in (a, b)$ and hence $f''(c) \geq 0, \forall c \in (a, b)$. i.e,

$$\frac{f'(x_2) - f'(x_1)}{x_2 - x_1} \geq 0 \tag{2.10}$$

Since, $x_2 \geq x_1$ i.e, $x_2 - x_1 \geq 0$ therefore for inequality holds (2.10), $f'(x_2) - f'(x_1) \geq 0$ and hence $f'(x_2) \geq f'(x_1)$. By definition of increasing function, $f'(x)$ is increasing.

(ii)⇒(iii): Let f' is monotonically increasing and f'' exists then

$$f''(x) = \lim_{h \to 0} \frac{f'(x + h)}{f'(x)}$$

for $h > 0, x < x+h$. since f' is increasing then $f'(x) \leq f'(x+h)$ i.e, $f'(x+h) - f'(x) \geq 0$. Hence $f''(x) \geq 0$.

(i)$\Rightarrow$ (iii) Since f is convex then by above theorem (2.1.5), f' is monotonically increasing on (a,b). Similarly,(iii)$\Rightarrow$(i) Since, f' is increasing on $(a, b)1$ and f'' exists.

(i)$\Rightarrow$(iv): Assume that f is convex. Let $y > x_0$ and $y = x_0 + h$.

$$f'(x) = \lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h} \leq \frac{f(y) - f(x_0)}{y - x_0}$$

i.e, $f'(x_0)(y - x_0) \leq f(y) - f(x_0)$ since$(y - x_0) > 0$
Again, Let $y < x_0$

$$f'(x) = \lim_{h \Rightarrow 0^-} \frac{f(x_0 + h) - f(x_0)}{h} \geq \frac{f(y) - f(x_0)}{y - x_0}$$

i.e, $f'(x_0)(y - x_0) \leq f(y) - f(x_0)$. since $(y - x_0) > 0$.

(iv)$\Leftarrow$(i): We have given that

$$f'(x)(y - x_0) \leq f(y) - f(x_0) \tag{2.11}$$

holds for all $x_0, y \in \mathbb{E}$.
Now s ince equation (2.11) holds. Let $x, y \in \mathbb{E}$,take $z = \lambda x + (1 - \lambda)y$ then for x, z we can write ,

$$f(x) \geq f(z) + f'(z)(x - z) \tag{2.12}$$

$$\text{and} \quad \text{for} \quad \text{y, z}$$

$$f(y) \geq f(z) + f'(z)(y - z) \tag{2.13}$$

Multiply by λ in equation(2.12)and $(1 - \lambda)$ in equation(2.13) then adding both we get,

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) &\geq f(z) + f'(z)(\lambda x + (1 - \lambda)y - z) \\ &= f(z) + 0 = f(z) \end{aligned}$$

Hence $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ and hence f is convex. $\square$

A special form of convex function is the following:

Definition 2.1.6 (**Midpoint Convex Function in $\mathbb{R}$**)**.** A function $f : \mathbb{E} \to \mathbb{R}$ is a midpoint convex if for any $x, y \in \mathbb{R}$,

$$f\left(\frac{x + y}{2}\right) \leq \frac{f(x) + f(y)}{2}. \tag{2.14}$$

Example 2.1.7. Consider the function $f : \mathbb{R} \to (0, \infty)$ given by $f(x) = \exp(x)$. Now we check the condition of midpoint convex. Let $x, y \in \mathbb{R}$

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$
$$or, \quad \exp\left(\frac{x+y}{2}\right) \leq \frac{\exp x + \exp y}{2} \tag{2.15}$$

Since, $2ab \leq a^2 + b^2$. So, if we take $a = \exp\left(\frac{x}{2}\right)$ and $b = \exp\left(\frac{y}{2}\right)$ then we get, $\exp\left(\frac{x+y}{2}\right) \leq \frac{\exp(x)+\exp(y)}{2}$. Hence equation (2.15) is true.
Thus, f is midpoint convex and it is obviously continuous. So, $\exp(x)$ is convex.

Theorem 2.1.8. Let $f : \mathbb{E} \to \mathbb{R}$ be a continuous function. Then the following are equivalent,

(i) f is convex

(ii) f is midpoint convex.

Proof. (i)⇒(ii) This Proof is trivial when we take $\lambda = \frac{1}{2}$.
(ii)⇒(i) First we shall prove that

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

for all "dyadic rational" numbers, i.e., all numbers of the form $\lambda = \frac{k}{2^n}$, where k is a non negative integer not less than 2^n. We do this by induction on n. The case $n = 0$ is trivial (Since $\lambda = 0$ or $\lambda = 1$). In the case $n = 1$ we have $\lambda = 0$ or $\lambda = 1$ or $\lambda = \frac{1}{2}$. The first two cases are again trivial, and the third is precisely the hypothesis of the theorem. Suppose the result is proved for $n \leq r$, and consider $\lambda = \frac{k}{2^{r+1}}$. If k is even, say $k = 2l$, then $\frac{k}{2^{r+1}} = \frac{l}{2^r}$, and we can appeal to the induction hypothesis. Now suppose k is odd. Then $1 \leq k \leq 2^{r+1} - 1$, so the numbers $l = \frac{k-1}{2}$ and $m = \frac{k+1}{2}$ are integers with $0 \leq l < m \leq 2^r$. We can now write

$$\lambda = \frac{s+t}{2},$$

where $s = \frac{k-1}{2^{r+1}} = \frac{l}{2^r}$ and $t = \frac{k+1}{2^{r+1}} = \frac{m}{2^r}$. We then have

$$\lambda x + (1-\lambda)y = \frac{[sx + (1-s)y] + [tx + (1-t)y]}{2}.$$

Hence by the hypothesis of he theorem and the induction hypothesis we have

$$\begin{aligned} f(\lambda x+(1-\lambda)y) &\leq \frac{f(sx+(1-s)y)+(f(tx+(1-t)y))}{2} \\ &\leq \frac{sf(x)+(1-s)f(y)+(1-t)f(y))}{2} \\ &= \left(\frac{(s+t)}{2}\right)f(x)+\left(1-\frac{s+t}{2}\right)f(y) \\ &= \lambda f(x)+(1-\lambda)f(y). \end{aligned}$$

This completes the induction. Now for each fixed x and y both sides of the inequality

$$f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$$

are continuous functions of λ. Hence the set on which this inequality holds (the inverse mage of the closed set $[0,\infty)$ under the mapping $\lambda \mapsto \lambda f(x)+(1-\lambda)f(y)-f(\lambda x+(1-\lambda)y)$ is closed set. Since it contains all the points $\frac{k}{2^n}$, $0 \leq k \leq n, n = 1, 2, ...$ It must contain the closure of this set of points, i.e, it must contain all of $[0.1]$. Thus f is convex. $\square$

Recall 2.1.9. The mean value theorem of differentiation, contained with the to statements of the previous proposition, show that a real differentiable function f is convex on (a,b). i.e, $f'(s) \leq f'(t)$ if $0 < s < t < b$. Consequently, in a twice differentiable function f on (a,b), we have the useful characterization of a convex function:
If $f \in C^2(a,b)$, f is convex if and only if $f'' \geq 0$ on (a,b).

Theorem 2.1.10 (Jensen's Inequality)**.** Let (X, S, μ) be a probability measure space. Let

(i) ϕ be convex on $-\infty < a < b < \infty$;

(ii) $f \in L'(X,\mu)$

(iii) $f(X) \subset (a,b)$. Then

$$\phi\left(\int_X f d\mu\right) \leq \int_X (\phi \circ f) d\mu$$

Proof. If $a < z < b$ then it follows from the proposition (2.1.3) that the right hand derivative $D^+\phi(z) = \lim_{t\downarrow z} \frac{\phi(t)+\phi(z)}{t-z}$ and the left-hand derivative $D^+(z) = \lim_{t\uparrow z} \frac{\phi(t)+\phi(z)}{t-z}$ exists and

$$A := D^-\phi(z) \leq D^+\phi(z) := B$$

. Hence if $A \leq m \leq B$ and $a < t < b$, it follows that

$$\phi(t) \geq \phi(z) + m(t-z),$$

i.e, the graph of any supporting line lies below the graph of ϕ. if $z = \int_X f d\mu, a < x < b$,

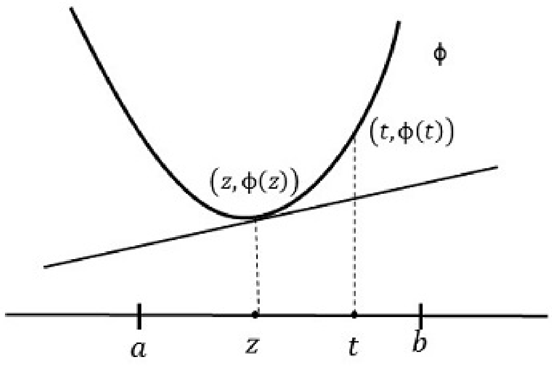


FIGURE 2.3

and $t = f(x)$, then $t, z \subset (a, b)$ and

$$m(f(x) - z) + \phi(x) \leq \phi[f(x)].$$

ϕ continuous and f is increasing means that $\phi \circ f$ is measurable.Now integrating both side, the above inequality gives the sum as $\mu(X) = 1$, the inequality in question is follows. □

Example 2.1.11. Take $\phi(x) = \exp(x)$. then the inequality becomes

$$\exp\left(\int_X f d\mu\right) \leq \int_X \exp(f) d\mu \tag{2.16}$$

If X is a finite set consisting of points $p_1, ..., p_n$ and if

$$\mu(p_i) = \frac{1}{n}, f(p_i) = x_i \quad (1 \leq i \leq n),$$

the above inequality becomes

$$\exp\{\frac{1}{n}(x_1 + x_2 + ... + x_n)\} \leq \frac{1}{n}(\exp(x_1) + \exp(x_2) + ... + \exp(x_n))$$

for real x_i. Putting $y_i = \exp(x_i)$, we get

$$(y_1 y_2 ... y_n)^{\frac{1}{n}} \leq \frac{1}{n}(y_1 + y_2 + ... + y_n) \tag{2.17}$$

the familiar inequality $G.M. \leq A.M.$. Going back from this to (2.16), it should become clear why the left and right sides of the inequality

$$\exp\{\int_X \log g d\mu\} \leq \int_X g d\mu$$

are often called the geometric and arithmetic mean, respectively for the positive function g. If we take $\mu(p_i) = \alpha_i > 0$ where $\sum \alpha_i = 1$, we obtain

$$y_1^{\alpha_1}, ...y_n^{\alpha_n} \leq \alpha_1 y_1 + \alpha_2 y_2 + ...\alpha_n y_n$$

in place of (2.17). These are just a few samples of the wealth of information contained in Jensen's Inequality.

2.2 A few basic results of Convex function on $\mathbb{R}^n$

In this section we discuss the generalized form of convex function in n-dimensional space and assume that $\mathbb{E}$ be convex subset of $\mathbb{R}^n$.

Definition 2.2.1 (Convex Function in $\mathbb{R}^n$)**.** The function $f : \mathbb{R}^n \to \mathbb{R}$ is called convex function if for all $x, y \in \mathbb{E}$ and $\lambda \in [0, 1]$ following condition holds:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda f(y)) \tag{2.18}$$

Example 2.2.2. Following are examples of convex functions in $\mathbb{R}^n$

(i) If $f : \mathbb{R}^n \to \mathbb{R}$ and given by $f(x) = \|x\|^p, \quad 1 \leq p < \infty$ then f is a convex function. In particular, let $\|.\|$ be any norm on $\mathbb{R}^n$.Then $\|.\|$ is convex function.

(ii) Let X be a norm linear space and Y be a closed subspace of X and $f : X \to \mathbb{R} \geq 0$ be given by $f(x) = d(x, Y) = \inf_{y \in Y} \|x - y\|$. Then f is convex.

(iii) Let X be a norm linear space and Y be a closed subspace of X and $\phi : X \to \mathbb{R} \geq 0$ be given by $\phi(x) = \sup_{z \in Y} \|x - z\|$ Then ϕ is convex function.

Explaination:

(i) Consider two functions $h : [0, \infty) \to [0, \infty)$ and $g : \mathbb{R}^n \to [0, \infty)$ be defined as $h(x) = x^p$ and $g(x) = \|x\|$, where $p \geq 1$. Then $(h \circ g)(x) = h(g(x)) = \|x\|^p = f(x)$. By triangular inequality g is convex. h is non-decreasing and contiuous and differentiable in $[o, \infty)$. Then $h'(x), h''(x) \geq 0$.

(ii) Let $f(x) = d(x, y)$. Then $f(\lambda x + (1 - \lambda)z) = d(\lambda x + (1 - \lambda)z, y)$
$= \inf_{y \in Y} \|\lambda x + (1 - \lambda)z - y\| = \inf_{y \in Y} \|\lambda(x - y) + (1 - \lambda)(z - y)\|$
$\leq |\lambda| \inf_{y \in Y}(\|x - y\|) + |(1 - \lambda)| \inf_{y \in Y}(|\|z - y\| = \lambda f(x) + 1 - \lambda)f(z)$.
Hence f(x) is convex.

(iii) Let $\phi(x) = d(x, Y) = \sup_{z\in Y} \|x - z\|$. Then $\phi(\lambda x + (1-\lambda)z) = d(\lambda x + (1-\lambda)z, y)$ $= \sup_{z\in Y} \|\lambda x + (1-\lambda)z - y\| = \sup_{z\in Y} \|\lambda(x-y) + (1-\lambda)(z-y)\|$ $\leq |\lambda| \sup_{z\in Y}(\|x-y\|) + |(1-\lambda)| \sup_{z\in Y}(\|z-y\|) = \lambda\phi(x) + 1 - \lambda)\phi(z)$. Hence ϕ is convex.

Theorem 2.2.3. Let $f : \mathbb{E} \to \mathbb{R}$. Then f is convex if and only if $epi(f)$ is convex in $\mathbb{E} \times \mathbb{R}$.

Proof. **(**$\Leftarrow$**)** Assume that $epi(f)$ is convex. Let $x_1, x_2 \in \mathbb{R}$ and $\lambda \in [0,1]$. Then the points $(x_1, f(x_1))$ and $(x_2, f(x_2)$ are in the $epi(f)$. Since, the $epi(f)$ is convex.
Let the point $(\hat{x}, \hat{y}) = (\lambda x_1 + (1-\lambda)x_2, \lambda f(x_1) + (1-\lambda)f(x_2)$ is in the $epi(f)$. i.e, $(\hat{x}) \in epi(f)$. Then by def of $epi(f)$

$$\begin{aligned} & \hat{y} \geqslant f(\hat{x}) \\ i.e, \quad & f(\hat{x}) \leq \hat{y} \\ or, \quad & f(\lambda x_1 + (1-\lambda)x_2) \leq f(x_1) + (1-\lambda)f(x_2) \end{aligned}$$

so, $f(\lambda x_1 + (1-\lambda))$. Hence $epi(f)$ is convex.

($\Rightarrow$**)** We need to show that if $f$is convex then $epi(f)$ must be convex.
So, let (x_1, x_2) and (x_2, y_2) be in the epigraph of f(x). suppose $(\bar{x}, \bar{y}) = (\lambda x_1 + (1-\lambda)y_2)$,where $\lambda \in [0,1]$.then we have,

$$\begin{aligned} \bar{y} &= \lambda x_1 + (1-\lambda)Y_2 \\ &\geqslant \lambda y_1 + (1-\lambda)y_2 \\ &\geqslant f(\lambda x_1 + (1-\lambda)f(x_2) = f(x) \end{aligned}$$

i.e, $\quad f(\bar{x}) \leq \bar{y} = \lambda y_1 + (1-\lambda)y_2$. Hence, $epi(f)$ is convex. □

Theorem 2.2.4. Let $f : E \to \mathbb{R}$ is twice differentiable function.Then the following are equivalent,

(i) f is convex.

(ii) $\nabla f(\vec{x_0})(\vec{y} - \vec{x_0}) \leq f(\vec{y}) - f(\vec{x_0}) \quad \forall \vec{x_0}, \vec{y} \in \mathbb{R}^n$

(iii) $\vec{y}^T \nabla^2 f(\vec{x_0})\vec{y} \geq 0, \quad \forall \vec{y} \in E$

Proof. (i)$\Rightarrow$(ii): Define $g : [0,1] \to \mathbb{R}$ by $\quad g(t) = f(\vec{x_0} + t(\vec{y} - \vec{x_0}))$.
Claim: g is convex and differentiable.

Proof of the claim: Let $t_1, t_2 \in [0,1]$ then

$$\begin{aligned} g(\lambda t_1 + (1-\lambda t_2)) &= f[\vec{x_0} + (\lambda t_1 + (1-\lambda)t_2)(\vec{y} - \vec{x_0})] \\ &= f(\lambda\vec{x_0} + \lambda t_1(\vec{y} - \vec{x_0}) + \vec{x_0} - \lambda\vec{x_0} + (1-\lambda)t_2(\vec{y} - \vec{x_0})) \\ &= f(\lambda(\vec{x_0} + t_1(\vec{y} - \vec{x_0})) + (1-\lambda)(\vec{x_0} + t_2(\vec{y} - \vec{x_0}))) \end{aligned}$$

Since f convex, therefore

$$\begin{aligned} &\leq \lambda(f(x_0 + t_1(\vec{y}) - \vec{x_0})) + (1-\lambda)f(\vec{x_0} + t_2(\vec{y} - \vec{x_0}))) \\ &= \lambda g(t_1) + (1-\lambda)g(t_2) \end{aligned}$$

Hence g is convex. Define $\phi : [0,1] \to \mathbb{R}^n$ by $\phi(t) = \vec{x_0} + t(\vec{y} + \vec{x_0})$ which is continuous and differentiable. So, $g : [0,1] \to \mathbb{R}$ which is given by $g(t) = (f \circ \phi)(t)$ is also continuous and differentiable. Since $g(t) = f(\vec{x_0} + t(\vec{y} - \vec{x_0}))$ then $g'(t) = \nabla f(\vec{x_0}) + t(\vec{y})$.
Now g is convex and $g : [0,1] \to \mathbb{R}$. Then by theorem (2.1.5), $g'(x_0)(y - x_0) \leq g(y) - g(x_0), \forall x_0, y \in [0,1]$. Take $x_0 = 0, y = 1$ we get, $g'(x_0)(1-0) \leq g(1) - g(0)$. Hence $\nabla f(\vec{x_0})(\vec{y} - \vec{x_0}) \leq f(\vec{y}) - f(\vec{x_0})$ and hence

$$\nabla f(\vec{x_0})(\vec{y} - \vec{x_0}) \leq f(\vec{y}) - f(\vec{x_0}) \tag{2.19}$$

(ii)$\Leftarrow$(i): Given that equation(2.19) holds for all $\vec{x_0}, \vec{y} \in \mathbb{R}^n$. Consider $\vec{x_0} = \lambda\vec{x_0} + (1-\lambda)\vec{y})$
Now, for $\vec{x_0}, \vec{x} \in \mathbb{R}^n$

$$f(\vec{x}) \leq f(\vec{x_0}) + \nabla f(\vec{x_0})(\vec{x} - \vec{x_0}) \tag{2.20}$$

Again for $\vec{x_0}, \vec{y}$

$$f(\vec{y}) \leq f(\vec{x_0}) + \nabla f(\vec{x_0})(\vec{y} - \vec{x_0}) \tag{2.21}$$

Multiplying (2.20) by λ and (2.21) by $(1-\lambda)$ we get,

$$\lambda f(\vec{x_0}) + (1-\lambda)f(\vec{y}) \geq \lambda f(\vec{x_0} + \lambda\nabla f(\vec{x_0})(\vec{x} - \vec{x_0})$$

i.e, $\lambda f(\vec{x}) + (1-\lambda)f(\vec{y}) \geq f(\vec{x_0}) + \nabla f(\vec{x_0})\,(\lambda(\vec{x} - \vec{x_0}) + (1-\lambda)(\vec{y} - \vec{x_0}))$
i.e, $\lambda f(\vec{x}) + (1-\lambda)f(\vec{y}) \geq f(\vec{x_0}) + \nabla f(\vec{x_0})\,(\lambda\vec{x} - (1-\lambda)(\vec{y} - \vec{x_0}))$

(i) $\Leftrightarrow$(iii): Define $g : [0,1] \to \mathbb{R}$ by $g(t) = f(\vec{x_0} + t(\vec{y} - \vec{x_0}))$. Clearly, g is convex and differentiable. $g'(t) = \nabla f(\vec{x_0} + t(\vec{y} - \vec{x_0}))(\vec{y} - \vec{x_0})$
Now f is convex if and only if g is convex. g is convex if and only if $g'' \leq 0$ Thus, it suffices to show that $g'' \leq 0$
Claim: $g'(t)$ is differentiable. Since, f and g is twice differentiable function. So, $g'(t) =$

$\nabla f(\vec{x_0} + t(\vec{y} - \vec{x_0}))$

$$
\begin{aligned}
g''(t) =& \frac{d}{dt} g'(t) = \lim_{h \to 0} \frac{g'(t+h) - g'(t)}{h} \\
=& \frac{\lim_{h\to 0} \left(\frac{df}{\delta x_i} (\vec{x_0} + (t+h)(\vec{y} - \vec{x_0})) \right)_{i=1}^{n} - \lim_{h\to 0} \left(\frac{df}{\delta x_i} (\vec{x_0} + (t)(\vec{y} - \vec{x_0})) \right)_{i=1}^{n}}{h} . (\vec{y} - \vec{x_0}) \\
=& (\vec{y} - \vec{x_0}) \lim_{h \to 0} \left(\frac{\frac{df}{\delta x_i} (\vec{x_0} + (t+h))(\vec{y} - \vec{x_0})) - \frac{df}{\delta x_i} (\vec{x_0} + (t)(\vec{y} - \vec{x_0}))}{h} \right)_{i=1}^{n} \\
=& (\vec{y} - \vec{x_0})^T \left(\frac{\delta^2 f}{\delta x_i \delta x_j} (\vec{x} + t(\vec{y} - x \vec{-} 0) \right) (\vec{y} - \vec{x_0}) \geq 0
\end{aligned}
$$

□

Proposition 2.2.5. Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex on E and E_M is defined by $E_M = \{x \in E : f(x) \leq M\}$ then E_M is a convex set for all M.

Proof. Let $x_1, x_3 \in E_M$ and $t \in [0, 1]$.Then
$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) \leq tM + (1-t)M = M$
Hence $f(tx_1 + (1-t)x_2) \leq M$ and hence $tx_1 + (1-t)x_2 \in E_M$

□

Theorem 2.2.6. Let $\mathbb{E}$ be an open convex set and f is convex on E. Then f is continuous on E.

Proof. Let $x_0 \in \mathbb{E}$ and $d = \inf\{\|x_0 - p\| : p \in \delta\mathbb{E}\}$, distance from x_0 to the boundary of $\mathbb{E}(d = \infty$ if $\mathbb{E} = \mathbb{R}^n)$. Let C be a n-cube with center x_0 and side length 2δ where $\delta < \frac{d}{\sqrt{n}}$.(See the following picture). This ensure that C lies inside $\mathbb{E}$ and the ball around x_0 of radius δ lies inside C. The set of vertices V of C is finite and and $C = \text{Conv}(V)$. Let

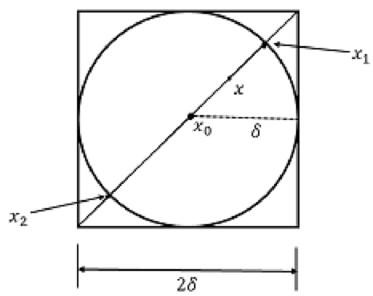


FIGURE 2.4

$M = \{Maxf(x) : x \in V\}$, by Proposition (2.2.5) we can say that $\mathbb{E}_M = \{x : |f(x)| < M\}$ is convex. Clearly see that $V \subseteq \mathbb{E}_M$. We need to check that $C \subset \mathbb{E}_M$.
To verify this, let $y \in C$ and $t \in [0,1]$. Here $y = ta + (1-t)b$ for all $a, b \in V$ which implies

$$f(y) = f(ta + (1-t)b) = tf(a) = (1-t)f(b) = tM + (1-t)M = M$$

Therefore $f(y) \leq M$ i.e, $y \in E_M$ hence $C \subset E_M$. Let X be any point such that $0 \leq \|x - x_0\| < \delta$ and line segment through x_0 and x intersect the ball at points x_1 and x_2. We write x and x_0 as linear combination of x_1, x_0 and x_2, x respectively. If $t = \dfrac{\|x - x_0\|}{\delta}$ then

$$x = tx_1 + (1-t)x_0$$

$$x_0 = \frac{\|x - x_0\|}{\delta + \|x - x_0\|} + \frac{\delta}{\delta + \|x - x_0\|}$$

Since f is convex, $f(x) \leq tf(x) + (1-t)f(x_0)$

$$f(x) - f(x_0) \leq tf(x_1) - tf(x_0) = t[f(x_1) - f(x_0)] = t[M - f(x_0)] \tag{2.22}$$

$$f(x_0) \leq \frac{t}{1+t} f(x_2) + \frac{1}{1+t} f(x)$$

$$(1+t)f(x_0) \leq tf(x_2) + f(x)$$

$$f(x_0) - f(x) \leq tf(x_2) - tf(x_0)$$

$$f(x) - f(x_0) \geqslant -t[f(x_2) - f(x_0)] = -t[M - f(x_0)] \tag{2.23}$$

$\mathrm{from}(2.22)\mathrm{and}(2.23); \mathrm{weget},$

$$-t[M - f(x_0)] \leq f(x) - f(x_0) \leq t[M - f(x_0)]$$

$$|f(x) - f(x_0)| \leq \|x - x_0\| [\frac{[M - f(x_0)]}{\delta}] \tag{2.24}$$

which shows that f satisfies a Lipschitz Condition of order 1 and is therefore continuous. □

Remark 2.2.7. It should be noted that the above result may fail if K is not open. For example, take $K = [0,1]$ and

$$f(x) = \begin{cases} x & \text{if} \quad 0 < x \leq 1 \\ 1 & \text{if} \quad x = 0 \end{cases} .$$

Then f is convex but is discontinuous at 0.

2.3 A Characterization of the Gamma Function by means of convexity

The Gamma function is one of the most important special functions of classical mathematics. It plays vital role in the proof of Stirling's formula and it has profound connections with the Riemann Zeta function. In this section we explore some properties of Gamma function that relate to convexity properties. We discuss the definitions and some properties of the Gamma function before coming to the result.

Definition 2.3.1. The Gamma function, $\Gamma(x)$ for $x > 0$ is usually defined by

$$\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt. \tag{2.25}$$

Lemma 2.1. *The improper integral*

$$\int_0^\infty e^{-t}t^{x-1}dt \tag{2.26}$$

converges for all $x > 0$.

Proof. We can write improper integral (2.26) as

$$\int_0^\infty e^{-t}t^{x-1}dt = \int_0^1 e^{-t}t^{x-1}dt + \int_1^\infty e^{-t}t^{x-1}dt$$

In order to show the convergence of (2.26). We will show that convergence of

$$\int_0^1 e^{-t}t^{x-1}dt \tag{2.27}$$

and

$$\int_1^\infty e^{-t}t^{x-1}dt \tag{2.28}$$

For the convergence of(2.27), we observe that $0 < e^{-t}t^{x-1} \leq t^{x-1}$ holds true for all $t \in [0, 1]$. Therefore, for $\epsilon > 0$ sufficiently small, we have

$$\int_\epsilon^1 e^{-t}t^{x-1}dt \leq \int_0^1 t^{x-1}dt = \frac{t^x}{x}\bigg|_\epsilon^1 = \frac{1}{x} - \frac{\epsilon^x}{x}.$$

Consequently, for all $x > 0$, the integral (2.27) is convergent. To assure the convergence of (2.28) we observe that

$$e^{-t}t^{x-1} = \frac{1}{\sum_{k=0}^{\infty}\frac{r^k}{k!}}t^{x-1} \le \frac{1}{\frac{1}{\frac{t^n}{n!}}}t^{x-1} = \frac{n!}{t^{n-x+1}}$$

for all $n \in \mathbb{N}$ and $t \ge 1$. This shows us that

$$\int_1^A e^{-t}t^{x-1}dt \le n!\int_1^A \frac{1}{t^{n-x+1}}dt = n!\frac{t^{-n+x}}{x-n}\bigg|_1^A = \frac{n!}{x-n}\left(\frac{1}{A^{n-x}} - 1\right)$$

provided that A is sufficiently large and n is chosen such that $n \ge x+1$. Thus (2.28) is convergent. □

We observe that $\Gamma(x) > 0$ for all $x > 0$ and by using integration by parts we find that

$$\begin{aligned}
\Gamma(x+1) &= \int_0^\infty e^{-t}t^x dt = -\int_0^\infty t^x de^{-t} \\
&= -e^{-t}t^x\bigg|_0^\infty + \int_0^\infty e^{-t}dt^x \\
&= x\int_0^\infty e^{-t}t^{x-1}dt \\
&= x\Gamma(x) \quad , x > 0
\end{aligned}$$

Further we have,

$$\Gamma(1) = \int_0^\infty e^{-t}dt = -e^{-t}\bigg|_0^\infty = 1$$

Lemma 2.2. *The function $\Gamma(x)$ satisfies the functional equation*

$$\Gamma(x+1) = x\Gamma(x) \ , x > 0 \tag{2.29}$$

Moreover, by iteration for $x > 0$ and $n \in \mathbb{N}$. We get,

$$\Gamma(x+n) = x(x+1)(x+2)......(x+n-1)\Gamma(x) \tag{2.30}$$

$$= \prod(x+i-1)\Gamma(x)$$

$$\Gamma(n+1) = \left(\prod_{i=1}^n\right)\Gamma(1) = \prod_{i=1}^n i = n! \tag{2.31}$$

Hence $\Gamma(n+1) = n!$.

In other words, the Gamma function is interpreted as an extension of factorials.

Lemma 2.3. *The function $\Gamma(x)$ is differentiable for all $x > 0$ and we have*

$$\Gamma'(x) = \int_0^\infty e^{-t} ln(t) t^{x-1} dt \tag{2.32}$$

Proof. For $x > |h| > 0$, we use the formula

$$t^y = e^y ln(t), \quad y > 0, \ t > 0$$

$$\Gamma(x+h) = \int_0^\infty e^{-t} t^{x+h-1} dt = \int_0^\infty e^{-t+ln(t)(x+h-1)} dt$$

By Taylor's formula, we find $0 < v < 1$ such that

$$\Gamma(x+h) - \Gamma(x) = \int_0^\infty e^{-t} e^{-ln(t)} \left(e^{ln(t)(x+h)} - e^{ln(t)x}\right) dt$$

$$= \int_0^\infty e^{-t} e^{-ln(t)} \left(hln(t).t^x + \frac{1}{2} h^2 ln(t) t^{(x+\theta h)}\right) dt$$

$$= h \int_0^\infty e^{-t} ln(t) t^{x-1} dt + \frac{1}{2} h^2 \int_0^\infty e^{-t} \left(ln(t)\right)^2 t^{x+\theta h-1} dt$$

This give us the differentiation of Γ if the second interval is bounded.
Consider the following estimate

$$\left(ln(t)\right)^2 \leq t^2 \quad \text{for} \quad t \geq 1 \quad \text{and} \quad e^{-t} \leq 1 \quad \text{for } t \in [0,1]$$

$$\begin{aligned}
\int_0^\infty e^{-t}(ln(t))^2 t^{x+vh-1} dt &= \int_0^1 e^{-t}(ln(t))^2 t^{x+vh-1} dt + \int_1^\infty e^{-t}(ln(t))^2 t^{x+vh-1} dt \\
&\leq \int_0^\infty (ln(t))^2 t^{x+vh-1} dt + \int_1^\infty e^{-t} t^2 t^{x+vh-1} dt \\
&\leq \Gamma(2+x+vh) + \frac{2}{x+vh^3} < \infty
\end{aligned}$$

This provide us desire result. $\square$

An analogous proof can be give to show that Γ is infinitely often differentiable for all $x > 0$ and

$$\Gamma^{(k)}(x) = \int_0^\infty e^{-t} (ln(t))^k t^{x-1} dt, \ k \in \mathbb{N} \tag{2.33}$$

Definition 2.3.2 (Logarithmic convex). If $f : (0, \infty) \to (0, \infty)$ then the function f is said to be logarithmically convex. If $log\left(f(x)\right)$ is convex on $(0, \infty)$.

Lemma 2.4. *For $x > 0$*

$$\Gamma'(x)^2 < \Gamma(x)\Gamma''(x) \tag{2.34}$$

Equivalently,

$$\left(\frac{d}{dx}\right)^2 ln(\Gamma(x)) = \frac{\Gamma''(x)}{\Gamma(x)} - \left(\frac{\Gamma'(x)}{\Gamma(x)}\right) > 0 \tag{2.35}$$

That means, $x \mapsto ln\Gamma(x),\ x > 0$ is convex function, or, Γ is ***logarithmic Convex****.*

Proof. We start with

$$\begin{aligned}(\Gamma'(x))^2 &= \left(\int_0^\infty e^{-t} ln(t) t^{x-1} dt\right)^2 \\ &= \left(\int_0^\infty e^{-\frac{t}{2}} t^{\frac{x-1}{2}} .ln(t).e^{\frac{-t}{2}} t^{\frac{x-1}{2}} dt\right)^2\end{aligned}$$

The Cauchy-Schwartz inequality yields (not that equality cannot occur since the two functions are linearly independent):

$$\begin{aligned}(\Gamma'(x))^2 &< \int_0^\infty \left(e^{-\frac{t}{2}} t^{\frac{x-1}{2}}\right)^2 dt. \int_0^\infty \left(e^{-\frac{t}{2}} t^{\frac{x-1}{2}} ln(t)\right)^2 dt \\ &= \int_0^\infty e^{-t} t^{x-1} dt. \int_0^\infty e^{-t} (ln(t))^2 t^{x-1} dt = \Gamma\Gamma''\end{aligned} \tag{2.36}$$

Moreover, we find that the help of (2.36) that

$$\frac{d^2}{dx^2} ln(\Gamma(x)) = \frac{\Gamma''(x)}{\Gamma(x)} - \left(\frac{\Gamma'(x)}{\Gamma(x)}\right) > 0$$

which yields (2.35). Note that $ln\Gamma(x)$ is convex i.e, for $t \in [0, 1]$

$$\begin{aligned}ln\left(\Gamma(tx + (1-t)y)\right) &\le t ln\Gamma(x) + (1-t) ln\Gamma(x) \\ &= ln\Gamma^t(x) + ln\Gamma^{1-t}(y) \\ &= ln\left(\Gamma^t(x).\Gamma^{1-t}(y)\right)\end{aligned} \tag{2.37}$$

which is equivalent to $\Gamma\left(tx + (1-t)y\right) \le \Gamma'(x).\Gamma^{1-t}(y)$ with $x, y > 0$. $\square$

Define $(a)_k = a(a+1)...(a+k-1)$, where $(a)_0 = 1$ and $k \in \mathbb{N}$
Thus, For $n \in \mathbb{N}$, $(x)_{n+1} = x(x+1)...(x+n)$, where $(x)_0 = 1$.

Lemma 2.5. *For $x > 0$ and $n \in \mathbb{N}$,*

$$\int_0^1 (1-t)^n t^{z-1} dt = \frac{n!}{(x)_{n+1}} \tag{2.38}$$

where $(x)_{n+1} = x(x+1)...(x+n)$

Proof. We will prove above equation by induction. In order to prove (2.38) by induction. We first take $n = 0$ to obtain for $x > 0$

$$\int_0^1 t^{x-1} dt = \frac{t^x}{x}\bigg|_0^1 = \frac{1}{2} = \frac{0!}{(2)_1}$$

Now we suppose that (2.38) holds for $n = k$. Then we have

$$\begin{aligned}
\int_0^1 (1-t)^{k+1} t^{x-1} dt = \int_0^1 &= \int_0^1 (1-t)(i-t)^{k]t^{x-1}} dt \\
&= \int_0^1 (1-t)^k t^{x-1} dt - \int_0^1 (i-t)^k t^x dt \\
&= \frac{k!}{(x)_{k+1}} - \frac{k!}{(x+1)_{k+1}} \\
&= \frac{(k+1)!}{(x)_{K+2}}
\end{aligned}$$

which is (2.38) for $n = k+1$. This shows that (2.38) holds for all $n \in \mathbb{N}$ $\square$

Now, we let $t = \dfrac{u}{n}$ into (2.38) to find that

$$\frac{1}{n^x} \int_0^n \left(1 - \frac{u}{n}\right)^n u^{x-1} du = \frac{n!}{(x)_{n+1}}$$

$$\int_0^n \left(1 - \frac{u}{n}\right)^n u^{x-1} du = \frac{n! n^x}{(x)_{n=1}} \tag{2.39}$$

$$\lim_{n\to\infty} \left(1 - \frac{u}{n}\right)^n = e^{-u}$$

we conclude that from (2.38)

$$\Gamma(x) = \int_0^\infty e^{-u} u^{x-1} du = \lim_{n\to\infty} \frac{n! n^x}{(x)_{n+1}}$$

Lemma 2.6. *For $x > 0$ and $n \in \mathbb{N}$,*

$$\Gamma(x) = \lim_{n\to\infty} \frac{n! n^x}{x(x+1)....(x+n)} \tag{2.40}$$

Where $\Gamma(x+1) = x\Gamma(x)$

Note that above equation (2.40) is known as **Gauss's Formula**.
Finally, we have the following important properties of Gamma function:

1. $\Gamma(1) = 1$
2. $\Gamma(x+1) = x\Gamma(x)$
3. $log\Gamma(x)$ is convex.

Now we come to the result that "These three properties characterize the gamma function" According to

Theorem 2.3.3 (The Bohr-Mollerup Theorem). Let f be a function defined on $(0, \infty)$ such that $f(x) > 0$ for all $x > 0$. Suppose that f has the following properties:

(i) $\log f(x)$ is convex.

(ii) $f(x+1) = xf(x)$, for all $x > 0$ and

(iii) $f(1) = 1$.
Then $f(x) = \Gamma(x)$ for all $x > 0$.

Proof. By (i) and (ii), we have

$$f(x+n) = x(x+1)......(x+n-1)f(x) \tag{2.41}$$

for every non-negative integer n. Therefore it suffices to prove that $f(x) = \Gamma(x)$ for $0 < x < 1$. Then above equation will show that this identity holds for all $x > 0$. Also, note that we have $f(m) = (m-1)!$ for every integer $m \geq 1$.
Let $0 < x < 1$ and n be an integer larger than 2. By the convexity of $\log f(x)$ from proposition(2.1.3) we get,

$$\frac{\log f(n) - \log f(n-1)}{n-(n-1)} \leq \frac{\log f(x+n) - \log f(n)}{(x+n)-n} \leq \frac{\log f(n+1) - \log f(n)}{(n+1)-n}$$
$$\text{i.e, } \log(n-1)! + \log(n-2)! \leq \frac{\log f(x+1) - \log f(n-1)!}{x} \leq x\log(n)! - \log(n-1)!$$
$$\text{or, } x\log(n-1)^x + \log(n-1)! \leq \log f(x+n) \leq \log n^x + \log(n+1)!$$

By the monotonicity of the exponential function, we get

$$(n-1)^x(n-1)! \leq f(x+n) \leq n^x(n-1)!$$

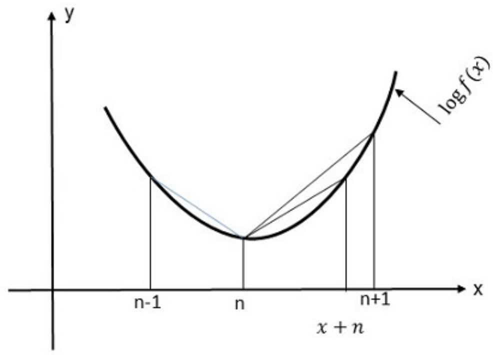


FIGURE 2.5

and (2.41) above shows that

$$\frac{(n-1)^x(n-1)!}{x(x+1)...(x+n-1)} \leq f(x) \leq \frac{n^x(n-1)!}{x(x+1)...(x+n-1)} = \frac{n^x n!}{x(x+1)...(x+n)}\left(\frac{x+n}{n}\right)$$

It is evident from the proof that one of the integers on the left hand side and Right hand side may be independent of one another and still preserves the inequality, Replacing n by $n+1$ on the left and keeping the right hand side unaltered, we get

$$\frac{n^x n!}{x(x+1)...(x+n)} \leq f(x) \leq \frac{n^x n!}{x(x+1)...(x+n)}\left[\frac{x+n}{n}\right]$$

for all $n > 2$ and $x \in (0,1)$. Letting $n \to \infty$ in the above inequality, we see from Gauss's formula that $f(x) = \Gamma(x), 0 \leq x \leq 1$ and that complete the proof. $\square$

Chapter 3

Convex functions on infinite dimensional spaces

This section provides an introduction to convex analysis, the proprieties of convex sets and functions.

3.1 Extended-valued convex functions

- The definition of convex(or concave) function that we have been using to date is not the most useful for the analysis of convex sets and functions. So far we allowed convex functions to be defined on convex subset of vector space.
- It is more useful to require that a convex(or concave) function be defined everywhere. We can do this by considering convex(or concave) to be extended real-valued function.
- If we take a function f that is convex in the sense of Definition (2.2.1), defined on the convex set K. we may extend it to entire vector space X by defining it to be ∞
- we start our discussion by taking the convexity of the epigraph to be the definition of convex function and allow convex functions to be extended real valued. First we discuss some useful definitions.

Definition 3.1.1 (Effective Domain). The set of points where a convex function does not assume the value ∞ is called its *effective domain*.

Definition 3.1.2 (**Proper Convex function).** A convex function is said to be *proper* if its effective domain is nonempty and additionally, it never assumes the value $-\infty$.

Definition 3.1.3 (**Extended convex functions).** An extended-real valued function $f : X \to R^* = [-\infty, \infty]$ on a vector space X is convex if its epigraph,

$$\text{epi}f = \{(x, \alpha) \in X \times \mathbb{R} : \alpha \geq f(x)\}$$

is convex subset of the vector space $X \times \mathbb{R}$.

The **effective domain** of a convex function f is the set $\{x \in X : f(x) < \infty\}$ and is denoted **dom** f.

★ Few simple facts

- Linear functions are both concave and convex.
- Any real-valued convex function defined on a non-empty convex subset K of X may be regarded as a proper convex function of all of X by putting $f(x) = \infty$ for $x \notin K$.
- The epigraph of extended real valued functions is a subset of $X \times \mathbb{R}$, not a subset of $X \times \mathbb{R}^*$.
- The effective domain of a convex function is convex set.
- The constant function $f = -\infty$ is convex but not proper(because its assumes the value $-\infty$).
- The constant function $g = \infty$ is convex but not proper(because its effective domain is empty).
- The function $g : \mathbb{R} \to \mathbb{R}^*$ defined by

  $$g(x) = \begin{cases} 0 & x = \pm 1 \\ \infty & |x| > 1 \\ -\infty & |x| < 1 \end{cases}$$

  is an example of non-trivial improper(not proper) convex function(Because it assumes the value $-\infty$ for $|x| < 1$).
- If a convex function is proper then its epigraph is a non-empty proper subset of $X \times \mathbb{R}$.

- Let f be an extended real-valued function on X. If f is finite at x the in fact x belongs to the interior of the effective domain of f.

- A convex function need not be finite at all points of continuity. The proper convex function f defined by
$$f(x) = \begin{cases} \dfrac{1}{x} & \text{for} \quad x > 0 \\ \infty & \text{for} \quad x \leq 0 \end{cases}$$
is continuous everywhere, even at $x = 0$ but not finite for $x \leq 0$.

Definition 3.1.4 (Lower semicontinuous convex functions). An extended real-valued proper convex function f on a topological vector space X is lower semicontinuous if and only if its epigraph is a closed (and convex) subset of $X \times \mathbb{R}$.

Corollary 3.1.5 (Closed convex sets)**.** In a locally convex space, if a convex set is not dense, the its closure is the intersection of all(topologically) closed half spaces that include it.

For a particular case, In a locally convex space X, Every proper closed convex subset of X is the intersection of all the closed half spaces that include it.

Recall 3.1.6. If f is proper convex function then epi(f) is a proper subset of $X \times \mathbb{R}$.

Thus, epi(f) is the intersection of all the closed half spaces that include it.

Lemma 3.1.7. Any non-vertical hyperplane in $X \times \mathbb{R}$ is the graph of some affine function on X. The graph of any affine function on X is some non-vertical hyperplane in $X \times \mathbb{R}$.

Proof. The non-vertical hyperplane

$$\{(x, \alpha) \in X \times \mathbb{R} : \langle (x*, \lambda), (x, \lambda) \rangle = c\}$$

where $\alpha \neq 0$, is the graph of the affine function,

$$g(x) = -\frac{1}{\lambda} x * (x) + \frac{c}{\lambda}$$

On the other hand, The graph of the affine function

$$x \mapsto x * (x) + c$$

is the non-vertical hyperplane. □

★ Few facts about separation of hyperplane

- The hyperplane $[f = \alpha]$ separates two sets A and B if either $A \subset [f \leq \alpha]$ and $B \subset [f \geq \alpha]$ or if $B \subset [f \leq \alpha]$ and $A \subset [f \geq \alpha]$.
- We say that hyperplane $[f = \alpha]$ properly separates A and B if it separates them and $A \cup B$ is not included in H.
- A hyperplane $[f = \alpha]$ strictly separates A and B if it separates them. In addition, $A \subset [f < \alpha]$ and $B \subset [f > \alpha]$ or vice-versa.
- We say that $[f = \alpha]$ strongly separates A and B if there is some $\epsilon > 0$ with $A \subset [f \leq \alpha]$ and $B \subset [f \geq \alpha + \epsilon]$ or vice-versa.
- We may also say that the linear functional f itself separates the sets when some hyperplane $[f = \alpha]$ separates them.

★ Basic topological separating hyperplane theorem

The following theorems are basic separating hyperplane theorems, which we will use many times to prove other theorems.

Theorem 3.1.8 (Basic Separating Hyperplane Theorem). Two nonempty disjoint convex subsets of a vector space can be properly separated by a nonzero linear functional,provided one of them has an internal point.

Theorem 3.1.9 (Interior Separating Hyperplane Theorem). In any topological vector space if the interiors of a convex set A is nonempty and is disjoint from another nonempty convex set B, then $\bar{A}$ and $\bar{B}$ can be properly separated by a nonzero continuous linear functional.

Theorem 3.1.10 (Strong Separating Hyperplane Theorem). For disjoint nonempty convex subsets of a (not necessarily Hausdorff) locally convex space, if one is compact and the other closed, then there is a nonzero continuous linear functional strongly separating them.

Corollary 3.1.11. In a locally convex space, if K be a non-empty closed convex set and $z \in K$, then there exists a non-zero continuous linear functional strongly separating K and z.

Definition 3.1.12. Let K be a nonempty closed convex subset of the vector space X, and Let $f : K \to \mathbb{R}$. Define the extended real functions f and f on K by

$$\hat{f}(x) := \{\inf g : g \geq f \quad \text{and} \quad g \quad \text{is afffne and continuous}\}$$

$$\tilde{f}(x) := \{\inf g : g \leq f \quad \text{and} \quad g \quad \text{is afffne and continuous}\}$$

where $\sup \phi = -\infty$ and $\inf \phi = \infty$.
The function $\hat{f}$ is called the **concave envelope** and $\tilde{f}$ is called **convex envelope**.

Lemma 3.1.13. Let X be a locally convex Hausdorff space and let $f : X \to \mathbb{R}^*$ be a lower semicontinuous proper convex function. If x belongs to effective domain and $\alpha < f(x)$ then there exist a continuous affine function g satisfying $g(x) = \alpha$ and $g \ll f$ means $g(y) < f(y)$ for all $y \in X$.

Proof. Observe that epi(f) is a non empty closed convex proper subset of $X \times \mathbb{R}$. By construction $\alpha < f(x)$ that tells that $(x, \alpha) \notin$ epi(f). Then by the Separating Hyperplane theorem, There exist a non zero linear functional $(x', \lambda) \in X' \times \mathbb{R}$(dual of $X \times \mathbb{R}$) and $\epsilon > 0$ satisfying

$$x'(x) + \lambda\alpha + \epsilon < x'(y) + \lambda\beta \tag{3.1}$$

forall $(y, \beta) \in$ epi(f). Since this inequality holds for β arbitrarily large, we have $\lambda \geq 0$. Now, x belongs to effective domain we have $f(x) < \infty$.
At $(y, \beta) = (x, f(x))$ equation (3.1) gives us

$$x'(x) + \lambda\alpha + \epsilon < x'(x) + \lambda f(x)$$

$$\lambda\alpha + \epsilon < \lambda f(x)$$

If $\lambda = 0$ then $\epsilon < 0$ but $\epsilon > 0$. So, it contradict. Hence $\lambda \neq 0$.
Now dividing by $\lambda > 0$ in (3.1)

$$x'(x) - x'(y) + \lambda\alpha + \epsilon < \lambda\beta$$

$$x'(x - y) + \lambda\alpha + \epsilon < \lambda\beta$$

$$\langle x - y, x' \rangle + \lambda\alpha + \epsilon < \lambda\beta$$

$$\frac{\langle x - y, x' \rangle}{\lambda} + \alpha + \frac{\epsilon}{\lambda} < \beta$$

Now the function $g : X \to \mathbb{R}$ defined by

$$g(z) = \frac{\langle x - z, x' \rangle}{\lambda} + \alpha$$

satisfying $g(x) = \alpha$ and forall $y \in \text{dom}(f)$, $(y, f(y)) \in \text{epi}(f)$

$$g(y) + \epsilon < f(y)$$

Hence $g(y) \ll f(y)$.
For $y \notin \text{dom} f$, $f(y) = \infty$ and $g(y) < \infty$. So, $g(y) < \infty = f(y)$. □

Theorem 3.1.14. Let X be a locally convex Hausdorff space and Let $f : X \to \mathbb{R}^*$ be a lower semicontinuous proper convex function. Then for all x,

$$f(x) = \sup\{g(x) : g \ll f \text{ and } g \text{ is affine and continuous}\}$$

Consequently, $f = \tilde{f}$.

Proof. Fix x and let $\alpha \in \mathbb{R}$ satisfy $\alpha < f(x)$ (Since f is proper ,$f \neq -\infty$. So, such a real α exists.) It suffices to show that there is a continuous affine function g with $g \ll f$ and $g(x) \leq \alpha$.
There are two cases to consider
Case(I): x belongs to effective domain:
Since $x \in \text{dom}(f)$ and $\alpha < f(x)$, by lemma (3.1) our sufficient condition covered. Case(II): x does not belongs to effective domain:
We proceed same as previous lemma to show that there exists a linear functional (x', λ) in dual space $X' \times \mathbb{R}$ and $\epsilon > 0$ satisfying (3.1) with $\lambda \geq 0$.
However, we may not conclude that $\lambda > 0$. So suppose that $\lambda = 0$. Then equation (3.1) becomes

$$x'(x) + \epsilon < x'(y)$$

for every $y \in \text{dom}(f)$.
Define the affine function $h : X \to \mathbb{R}$ by

$$h(z) = \langle x - z, x' \rangle + \frac{\epsilon}{2}$$

and observe that $h(x) > 0$ and for $y \in \text{dom}(f)$

$$h(y) = \langle x - y, x' \rangle + \frac{\epsilon}{2}$$

For every $y \in \text{dom}(f)$, $x'(x) + \epsilon < x'(y)$. That means

$$x'(x - y) + \frac{\epsilon}{2} < -\frac{\epsilon}{2}$$

Since $\epsilon > 0$. So, we have $h(y) < 0$.
Choose some $\bar{y} \in \text{dom}(f)$ and use previous lemma to find an affine function $\bar{g}$ satisfying $\bar{g} \ll f$. Now consider affine function $g : X \to \mathbb{R}$ of the form

$$g(z) = \gamma h(z) + \bar{g}(z) \qquad \text{where } \gamma > 0$$

For $y \in \text{dom}(f)$, we have $h(y) < 0$ and $\bar{g}(y) < f(y)$. So $g(y) < f(y)$, for $y \in \text{dom}(f)$.
For $y \notin \text{dom}(f)$, $g(y) < \infty = f(y)$. Thus $g(y) < f(y)$. $\square$

Remark 3.1.15. The epigraph of a lower semicontinuous proper convex function is a proper closed convex subset of $X \times \mathbb{R}$. Therefore it is the intersection of all the closed half spaces that include it.

3.2 Support Points

One of the recurring themes of this section is the characterization of the maxima and minima of linear functions over non empty convex sets.

Definition 3.2.1. Let A be a non-empty subset of topological vector space X, and Let f be a non-zero continuous linear functional on X. If f attains either its maximum or its minimum over A at the point $x \in A$. We say that f **supports** A **at** x and x is a **support point** of A.

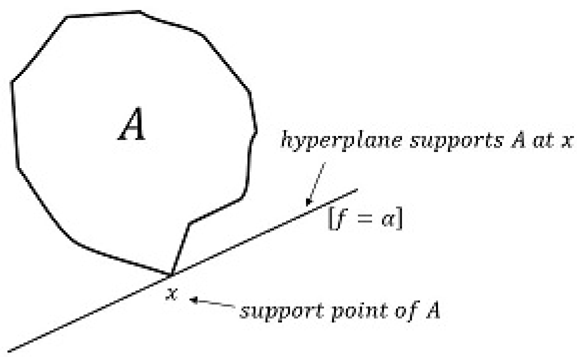


FIGURE 3.1: Support Point

- Taking $\alpha = f(x)$, we may also say that the hyperplane $[f = \alpha]$ supports A at x.
- If A is not wholly included in the hyperplane, we say that it is properly supported at x.
- We may also say that, the associated closed half space that includes A supports A at x.
  Here associated closed half space included A means we are taking half space $[f \leq \alpha]$ for a maximum at x and half space $[f \geq \alpha]$ for a minimum at x

To prove next theorem we need of following statements of these two theorems:

Theorem 3.2.2. If K is convex subset of a topological vector space then $0 < \alpha \leq 1$ then

$$\alpha K_0 + (1-\alpha)\overline{K}_0 \subset K_0$$

In particular, If $K_0 \neq \phi$ then

(i) The interior of K is dense in $\overline{K}$, i.e, $\overline{K_0} = \overline{K}$

(ii) The interior of $\overline{K}$ consides with the interior of K, i.e, $\overline{K}_0 = K_0$

Theorem 3.2.3. If A is a non-empty subset of topological vector space X and a non-zero linear functional f on X satisfies $f(x) \geq \alpha$ for all $x \in A$ then $f(x) > \alpha$ for all $x \in A_0$

Now we come to the theorem,

Theorem 3.2.4. Let K be a non-empty convex subset of a topological vector space X with non-empty interior.
If x is a boundary point of K that belongs to K, then K is a properly supported at x.

Proof. Let K°=Interior of K. Since K° is non-empty open convex set and $x \notin K^\circ$ (because x is a boundary point of K). Then by Interior separating Hyperplane theorem, There is a non-zero continuous linear functional f satisfying

$$f(x) \leq f(y) \text{ for all } y \in K^\circ$$

Since $K_0 \neq \phi$ by theorem (3.2.2), the interior of K is dense in $\overline{K_0} = \overline{K}$.
So, In fact $f(x) \leq f(y)$ for all $y \in K$ that means f supports C at the point x. Since $f(y) \geq f(x)$ for all $y \in K$ then by theorem (3.2.3), $f(y) > f(x)$ for all $y \in K_0$.
Hence the support is proper. □

Now we have an example which shows that above inclusion may fail for a convex set with empty interior.

Example 3.2.5. (**A boundary point which is not a support point**)
Consider the set l_1^+ of non-negative sequence in l_1, the Banach space of all summable sequences under the norm $\|.\|_1$-norm.

$$l_1^+ = \{(x_n) : \sum |x_n| < \infty, x_i \geq 0\} = \{(x_n) : \sum x_n < \infty, x_i \geq 0\}$$

It is clearly closed convex cone and its interior is empty.
To see this note that $\sum x_i < \infty$ then $x_i \to 0$ as $i \to \infty$ that means for each $\epsilon > 0$ and every $x = (x_1, x_2, ...) \in l_1^+$ there exists n_0 such that $x_{n_0} < \epsilon$.
Define $y = (y_1, y_2, ...)$ by $y_i = x_i$ for $i \neq n_0$ and $y_{n_0} = -\epsilon$
Now $\|x - y\| = \sum_{i=1}^{\infty} |x_i - y_i| = |x_{n_0} - y_{n_0}| < 2\epsilon$ that means $y \in B(x, \epsilon_0)$. Since ϵ is arbitrary, $B(x, \epsilon_0) \subseteq l_1^+$. But above assumption show that $y \notin l_1^+$ (because $y_{n_0} < 0$). This shows that l_1^+ has empty interior.
Thus, every point in l_1^+ is a boundary point. But no strictly positive sequence in l_1^+ is a support point.
To see this we make use of the fact that the dual space of l_1^+ is l_∞, the space of bounded

sequences. Let $x = (x_1, x_2, ...)$ be the element of l_1 such that $x_i > 0$ holds for each i, and suppose some nonzero $y = (y_1, y_2, ...) \in l_\infty$ satisfies

$$\sum_{i=1}^{\infty} y_i x_i \leq \sum_{i=1}^{\infty} y_i z_i$$

for all $z = (z_1, z_2, ...) \in l_1^+$. Letting $z_k = x_k + 1$ and $z_i = x_i$ for $i \neq k$ yeilds $y_k \geq 0$ for all k. Since y is nonzero, we must have $y_k > 0$ for some k, so $\sum_{i=1}^{\infty} y_i x_i > 0$.
But then $z = 0$ implies $0 = \sum_{i=1}^{\infty} y_i z_i \geq \sum_{i=1}^{\infty} y_i x_i > 0$, which is impossible. Thus x cannot be a support point.
On the other hand, if some $x \in l_1^+$ has $x_k = 0$ for some k, then the nonzero continuous linear functional $e_k \in l_{iy}$ satisfies

$$0 = e_k(x) \leq e_k(z)$$

for all $z \in l_+$. This means that e_k supports the set l_+ at x. Moreover, note that the collection of such x is norm dense in l_+.

3.3 Subgradients

Definition 3.3.1 (Subgradients). Given a dual pair $\langle X, X' \rangle$ and a convex function f on X. We say that $x' \in X'$ is a Subgradient of f at x, if it satisfies the following subgradient inequality

$$f(y) \geq f(x) + x'(y - x) \quad \text{forall } y \in X$$

- The set of subgradients at x is **subsdifferential of** f, denoted by $\partial f(x)$. i.e,

$$\partial f(x) = \text{set of subgradients at x.}$$

- It may be that $\partial f(x)$ is empty. But if $\partial f(x)$ is non-empty, we say that f is **subdifferentiable** at x

- For concave function f, If $x' \in X'$ satisfies the reverse subgradient inequality. i.e, If forall y, $f(y) \leq f(x) + x'(y - x)$ then we say that x' is a **supergradient** of f at x, and refer to the collection of them as **superdifferential**, also denoted $\partial f(x)$.

Lemma 3.3.2. A convex function f is minimized at x if and only if $0 \in \partial f(x)$

Proof. By taking $x' = 0$ in subgradient inequality

$$f(x) + x'(y - x) \leq f(y) \quad \text{forall } y \in X$$

it becomes $f(x) \leq f(y)$ forall $y \in X$. This shows that f minimized at x for $y \in X$. $\square$

- If f is a proper convex function. Then by considering $y \in \text{dom} f$, we see that the subgradient inequality can only be satisfied if $f(x) < \infty$, that is if $x \in \text{dom} f$.

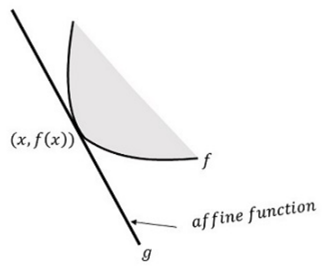
Another way to phrase the subgradients inequality is that
x' is a **subgradient** of f at $x \in \text{dom} f$ if f dominates the affine function

$$g(y) = x'(y) - x'(x) + f(x)$$

which agrees with f at x. That means

$$f(y) \geq g(y) \text{ for all } y \in X$$

and $f(y) = g(y)$ at $y = x$.



- The graph of an affine function is a non-vertical hyperplane in $X \times \mathbb{R}$. So, the subgradient inequality implies that this hyperplane supports the epigraph of f at $(x, f(x))$.

Lemma 3.3.3. The functional x' is a subgradient of the proper convex function f at $x \in \text{dom}(f)$ if and only if $(x, f(x))$ maximizes the linear functional $(x', -1)$ over $\text{epi}(f)$.

Proof. x' is subgradiant of the proper convex function f at $x \in \text{dom} f$

$$f(y) \geq f(x) + x'(y - x) \quad \text{for all } y \in X$$

$$\begin{aligned}
\alpha \geq f(y) &\geq f(x) + x'(y-x) \\
&= f(x) + x'(y) - x'(x) \\
x'(x) - f(x) &\geq x'(y) - \alpha \\
(x', -1)(x, f(x)) &\geq (x', -1)(y, \alpha) \quad \forall (y, \alpha) \in \text{epi}(f)
\end{aligned}$$

So, it shows that $(x', -1)$ maximizes at $(x, f(x))$ over epi(f). □

We now relate the subdifferential to the directional derivatives of f.

Lemma 3.3.4. Let f be a proper convex function, Let $x \in \text{dom}(f)$, $v \in X$ and $0 < \mu < \lambda$. Then the difference quotients satisfy

$$\frac{f(x+\mu v) - f(x)}{\mu} \leq \frac{f(x+\lambda v) - f(x)}{\lambda}$$

In particular,

$$\lim_{\lambda \to 0+} \frac{f(x+\lambda v) - f(x)}{\lambda} \text{ exists in } \mathbb{R}^*$$

Proof. The point $x + \mu v$ is the convex combination of $x + \lambda v$ and x

$$x + \mu v = \frac{\mu}{\lambda}(x + \lambda v) + \frac{\lambda - \mu}{\lambda} x$$

So, by convexity $f(x + \lambda v) \leq \dfrac{\mu}{\lambda} f(x + \lambda v) + \dfrac{\lambda - \mu}{\lambda} f(x)$

$$\text{Dividing by } \mu > 0, \ \frac{f(x+\lambda v)}{\mu} \leq \frac{f(x+\lambda v)}{\lambda} + \left(\frac{1}{\mu} - \frac{1}{\lambda}\right) f(x)$$

$$\frac{f(x+\mu v) - f(x)}{\mu} \leq \frac{f(x+\lambda v) - f(x)}{\lambda}$$

Thus we get the desired inequality. □

Define the **one-sided directional derivative** $d^+f(x) : X \to \mathbb{R}^*$ at x by

$$d^+f(x)(v) = \lim_{\lambda \to 0+} \frac{f(x+\lambda v) - f(x)}{\lambda}$$

Remark 3.3.5. If f is subdifferentiable at x (i.e, $\partial f(x) \neq \phi$) then this limit is finite.

Proof. Rewrite the subgradient inequality

$$f(y) \geq f(x) + x'(y - x)$$

Let $x + \lambda v = y$ then $y - x = \lambda v$, by putting this into subgradient inequality we get,

$$x'(\lambda v) \leq f(x + \lambda v) - f(x)$$

Since x' is a linear functional, $\lambda x'(v) \leq f(x + \lambda v) - f(x)$
This gives us $x'(v) \leq \dfrac{f(x + \lambda v) - f(x)}{\lambda}$ where $y = x + \lambda v$.
In this case, the difference quotient is bounded below by $x'(v)$ for any $x' \in \partial f(x)$.
so the limit is finite. □

We now show that $d^+ f(x)$ is a positively homogeneous convex function.

Theorem 3.3.6. Let f be a proper convex function on the tvs X. The directional derivative mapping $v \mapsto d^+ f(x)(v)$ from X into $\mathbb{R}^*$ satisfies the following properties:

(a) The function $d^+ f(x)$ is a positively homogeneous convex function (that is , sub-linear) and its effective domain is a convex cone.

(b) If f is continuous at $x \in \mathrm{dom} f$, then $v \mapsto d^+ f(x)(v)$ is continuous and finite valued.

Proof. It is easy to see that $v \to d^+ f(x)(v)$ is homogeneous as,

$$d^+ f(x)(v) = \lim_{\lambda \to 0+} \frac{f(x + \lambda v) - f(x)}{\lambda}$$

$$\because \frac{f(x + \lambda \alpha v) - f(x)}{\lambda} = \alpha \frac{f(x + \lambda \alpha v)}{\alpha \lambda}$$

So, $d^+ f(x)(\alpha v) = \alpha d^+ f(x)(v)$. This also shows that effective domain is cone.
To see this, dom($d^+ f(x)$)= $\{v \in X \mid d^+ f(x)(v) < \infty\}$

$$d^+ f(x)(\alpha v) = \alpha d^+ f(x)(v) < \infty$$

This show that $\alpha v \in \mathrm{dom}(d + f(x))$ for $\alpha > 0$.

For convexity, observe that for $u, v \in \mathrm{dom}(d^+ f(x))$

$$\begin{aligned}\frac{f(x+\lambda(\alpha u+(1-\alpha)v))-f(x)}{\lambda} &= \frac{f(x+\lambda\alpha u+\lambda v-\lambda\alpha v)-f(x)}{\lambda}\\ &= \frac{f(\alpha x+\alpha\lambda u+x+\lambda v-\alpha x-\lambda\alpha v)-f(x)}{\lambda}\\ &= \frac{f(\alpha(x+\lambda u)+(1-\alpha)(x+\lambda v))-f(x)}{\lambda}\\ &\leq \frac{\alpha f(x+\lambda u)+(1-\alpha)f(x+\lambda v))-f(x)}{\lambda}\\ &= \alpha\frac{f(x+\lambda u)-f(x)}{\lambda}+(1-\alpha)\frac{f(x+\lambda v)-f(x)}{\lambda}\end{aligned}$$

and letting $\alpha \to 0+$ yields $d+f(x)(\alpha u+(1-\alpha)v) \leq \alpha d+f(x)(u)+(1-\alpha)d+f(x)v$.

Lemma 3.3.7. Let $f : K \to \mathbb{R}$ be a convex function, where K is convex subset of a vector space. Let $x \in K$ and suppose z satisfies $x+z \in K$ and $x-z \in K$. Let $\delta \in [0,1]$ then,

$$|f(x+\delta z)-f(x)| \leq \delta \max\{f(x+z)-f(x), f(x-z)-f(x)\}|$$

Proof.

$$x+\delta z = (1-\delta)x+\delta(x+z)$$

So, $f(x+\delta z) \leq (1-\delta)f(x)+\delta f(x+z)$

$$f(x+\delta z)-f(x) \leq \delta[f(x+z)-f(x)] \tag{3.2}$$

and replacing z by $-z$ gives

$$f(x-\delta z)-f(x) \leq \delta[f(x-z)-f(x)] \tag{3.3}$$

Also, Since $x = \frac{1}{2}(x+\delta z)+\frac{1}{2}(x-\delta z)$
we have, $f(x) \leq \frac{1}{2}f(x+\delta z)+\frac{1}{2}f(x-\delta z)$
multiplying by 2, $2f(x) \leq f(x+\delta z)-f(x-\delta z)$

$$f(x)-f(x+\delta z) \leq f(x-\delta z)-f(x). \tag{3.4}$$

combining (3.3) and (3.4) we obtain,

$$f(x)-f(x+\delta z) \leq f(x-\delta z)-f(x) \leq \delta[f(x-z)-f(x)]. \tag{3.5}$$

This in conjunction with(3.2) yields the conclusion of the lemma. □

Now by lemma (3.3.7), we have

$$|f(x+\lambda u)-f(x)| \leq \delta \max\{f(x+u)-f(x), f(x-u)-f(x)\}|$$

for $0<\lambda\leq 1$. So let $\epsilon>0$ be given. If f is continuous at x, there exists an absorbing circled neighborhood. V of 0 such that $u\in V$ implies $|f(x+u)-f(x)|<\epsilon$. Thus we have

$$|d^+f(x)(v)| \leq \frac{|f(x+\lambda v)-f(x)|}{\lambda} \leq \max\{f(x+v)-f(x), f(x-v)-f(x)\}$$

That is , $d^+f(x)$ is bounded above on V. So, by Lemma (3.3.6), it is continuous.

□

Theorem 3.3.8. For are a proper convex function f and a point $x\in$dom$f$, the following are equivalent:

(1) x' is a subgradient of f at X.

(2) (x', -1) is maximized over epi$f$ at $(x, f(x))$.

(3) $x' < d^+f(x)$.

Proof. The equivalence of (1) and (2) is just Lemma (3.3.3).
To see the equivalence of (2) and (3), first note that any point y can be written as, $y=x+\lambda v$ with $v=y-x$ and $\lambda=1$.
So, (2) can be rewritten as

$$\langle (x',-1)(x,f(x))\rangle \geq \langle (x',-1),(y,f(y))\rangle$$

$$\langle (x',-1),(x,f(x))\rangle \geq \langle (x',-1),(x+\lambda v, f(x+\lambda))\rangle \tag{3.6}$$

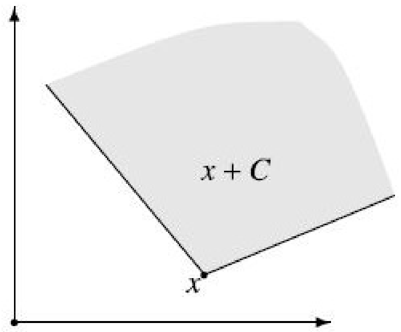
for all $y=x+\lambda v\in$dom$f$. Now note that (3.6) equivalent to the inequalities

$$x'(x)-f(x) \geq x'(x+\lambda v)-f(x+\lambda v) = x'(x)+x'(\lambda v)-f(x+\lambda v)$$

$$f(x+\lambda v)-f(x) \geq \lambda x'(v)$$

Thus, $x'(v) \leq \dfrac{f(x+\lambda v)-f(x)}{\lambda}$. In light of lemma (3.3.4) , this shows that (2) is equivalent to (3).

□



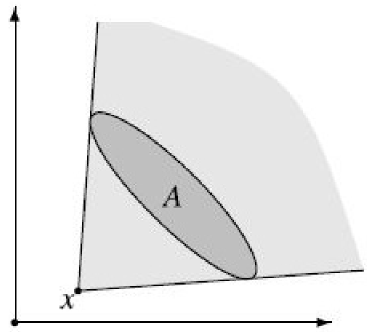


FIGURE 3.2: (i)A cone with vertex x (ii) A cone with vertex x generated by A

3.4 Supporting hyperplanes and Cone

This section refines the characterization of support points

Definition 3.4.1 (Cone). A cone is a nonempty subset of a vector space that is closed under multiplication by nonnegative scalars.

Definition 3.4.2 (Open Cone). An open cone is nonempty open subset of a topological vector space closed under multiplication by strictly positive scalars.

- An open cone contains the point zero only if it is the whole space.
- It is convenient to translate a cone or open cone around vector space. So we can say that, A nonempty subset of vector space is a cone with vertex x if it is of the form $x + C$, where C is a bona fide cone(with vertex 0).

Definition 3.4.3 (Convex Cone). A cone C is a convex cone if $\alpha x + \beta y \in C$, for any positive scalars α, β and any $x, y \in C$.

Example 3.4.4. (i) For a vector space X, the empty set, the space X, and any linear subspace of X are convex cones.

(ii) The set $\{x \in \mathbb{R}^2 | x_2 \geq 0, x_1 = 0\} \cup \{x \in \mathbb{R}^2 | x_1 \geq 0, x_2\}$ is a cone but not convex cone.

(iii) The set of positive semidefinite.

(iv) The intersection of two convex cones in the same vector space is again a convex cone, but their uinion may fail to be one.

(v) The norm cone
$$\{(x, r) \in \mathbb{R}^{d+1} | \|x\| \leq r\}$$
is a convex cone.

Theorem 3.4.5 (Klee)**.** In a locally convex space, A convex cone is supported at its vertex if the cone is not dense.

Proof. Let C be a convex cone with vertex x in a locally convex space. If C is supported at x, it lies in some closed half spaces and is thus not dense.
For the converse, assume that C is not dense. Without loss of generality we may assume that C is a cone with vertex 0. Now there exists some $x_0 \notin \overline{C}$. Since $\overline{C}$ is a closed convex set, the Separation Corollary (3.1.11) guarantees the existence of some nonzero continuous linear functional f satisfying $f(x_0) < f(x)$ for all $x \in C$.

□

Theorem 3.4.6. Let C be a convex subset of locally convex space and let x be a boundary point of K. If $x \in C$, then the following statements are equivalent:

(i) The set C is supported at x.

(ii) There is a non-dense convex cone with vertex X that includes C. or equivalently, The convex cone with vertex x generated by C is not dense.

(iii) There is an open convex cone K with vertex x such that $K \cap C = \phi$.

(iv) There exist a non-zero vector v and a neighborhood v of zero such that $x - \alpha v + z \in C$ with $\alpha > 0$ implies $z \notin \lambda v$.

Proof. (i) $\Rightarrow$ (ii) Any closed half space that supports C at x is a closed convex cone with vertex x that is not dense and includes C.
(ii) $\Rightarrow$ (iii) Let $\hat{K}$ be a non-dense convex cone with vertex X that includes C. By lemma (3.4.5), x is a point of support of $\hat{K}$. Now if f is a nonzero continuous linear functional attaining its maximum over $\hat{K}$ at x, then the open half space $K = [f > f(x)]$ is an open convex cone with vertex x that satisfies $K \cap C = \phi$.
(iii) $\Rightarrow$ (iv) Let K be an open convex cone with vertex 0 that satisfies

$$(x + K) \cap C = \phi$$

Fix a vector $w \in K$ and choose a neighborhood V of zero such that

$$w + V \subset K \text{ put } v = -w \neq 0.$$

Claim : v and V satisfy the desired properties. We have to show that $z \notin \alpha$ To see this, Assume that $x - \alpha v + z \in C$ with $\alpha > 0$. If $z = \alpha u$ for some $u \in V$ then

$$x - \alpha v + z = x - \alpha(-w) + \alpha u = x + \alpha w + \alpha u = x + \alpha(w + u) \in x + K$$

Since given that $x - \alpha v + z \in C$. Hence $x - \alpha v + z \in (x + K) \cap C$
That means tis contradict the assumption. Hence $z \in \alpha v$.
(iv)$\Rightarrow$ (i) We can assume that the neighborhood V of zero is open and convex.

The given condition grantees that the open convex cone K is generated by $-v + V$ with vertex zero, that is ,

$$K = \{\alpha(-v + w) : \alpha > 0 \text{ and } w \in V\}$$

satisfies $(x + K) \cap C = \phi$. Then, by Interior Separating Hyperplane Theorem (3.1.9), there exists a nonzero continuous linear functional f separating $x + K$ and C. That is, f satisfies

$$f(x + \alpha k) \leq f(y) \text{ for all } k \in K \quad \text{and all } y \in C.$$

Since K is a cone. we have

$$f(x + \alpha k) \leq f(y) \text{ for } \alpha > 0$$

and each $k \in K$, and by letting $\alpha \to 0^+$, we get $f(x) \leq f(y)$ for all $y \in C$. $\square$

The geometry of Theorem (3.4.6)is shown in figure (3.3). We can rephrase a separating

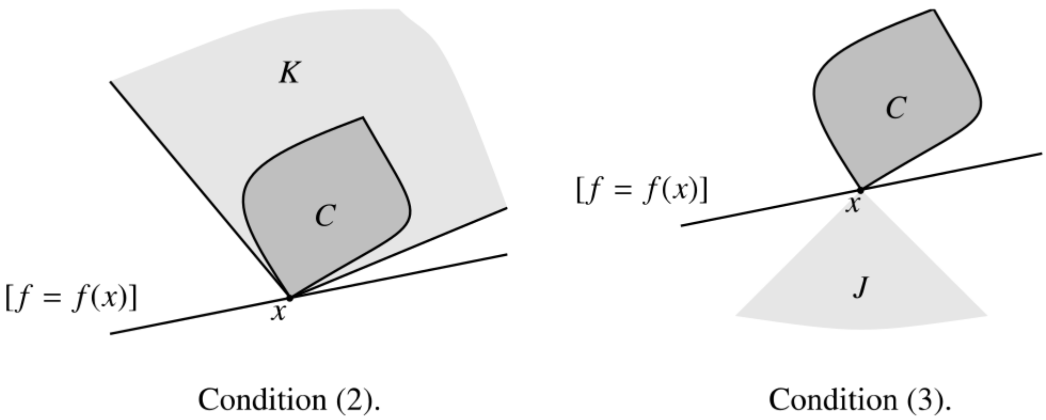


FIGURE 3.3

hyperplane theorem in terms of cones. Recall that the cone generated by S is the smallest cone that includes S and is thus $\{\alpha x : \alpha \geq 0\}$ and

3.5 The Bishop-Phelps Theorem

To understand the Bishop-Phelps Theorem we need a discussion of certain cones.

- Let X be a Banach Space. For $f \in X'$ with $\|f\| = 1$ and $0 < \partial < 1$, define

$$K(f, \partial) = \{x \in X : f(x) \geq \partial \|x\|\}.$$

  $K(f, \partial)$ is a closed convex pointed cone having a nonempty interior. To see this,

- In a Euclidean space, it can be described as the cone with major axis the half line determined by the unit vector f and having angle $w = arccos\ \delta$. See fig. (3.4).

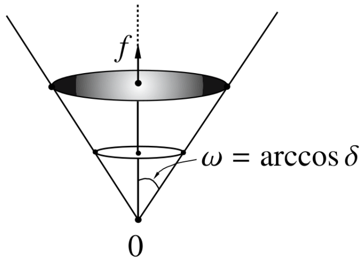


FIGURE 3.4

Now we will see that the cones $K(f, \partial)$ are related to the support points of convex sets. We need several properties that will be stated in terms of lemmas below.

Lemma 3.5.1. Let K be a nonempty convex subset of a Banach space and assume that $x_0 \in \partial K$. Then K is supported at x_0 if and only if there exists some cone of the form $K(f, \partial)$ satisfying $K \cap [x_0 + K(g, \partial)] = \{x_0\}$.

Proof. Assume that some continuous linear function g, $\|g\| = 1$, support C at c, that means $g(x) \leq g(c)$ for all $x \in C$. Then $C \cap [c + K(g, \delta)] = \{c\}$ for all $0 < \delta < 1$.
For the converse, assume that $C \cap [c + K(f, \delta)] = \{c\}$ for some continuous linear funcional of norm one and some $0 < \delta < 1$. If $K = c + K(f, \delta)$, then $C \cap K^0 = \phi$ and by Theorem (3.4.6) the set C is supported at c. □

Lemma 3.5.2 (Norm preserving extension)**.** A continuous linear functional defined on a subspace of a normed space can be extended to a continuous linear functional on the entire space while preserving its original norm.

Theorem 3.5.3 (Fundamental theorem of duality)**.** Let $f_1, f_2, f_3, ..., f_n$ be linear functionals on a vector space X. Then f lies in the span of $f_1, ..., f_n$ (that is, $f = \sum_{i=1}^{n} \lambda_i f_i$ for some scalrs $\lambda_1, ..., \lambda_n$) if and only if $\cap_{i=1}^{n} \ker f_i \subset \ker f$.

Lemma 3.5.4. Let f and g be norm one linear functionals on a Banach space X.
If for some $0 < \epsilon < 1$ we have $\|g|_{\ker f}\| \leq \epsilon$, that is, If the norm of g is restricted to the kernel of f is no more than ϵ, then either $\|f + g\| \leq 2\epsilon$ *or* $\|f - g\| \leq 2\epsilon$.

Proof. Let $\|g|_{\ker f}\| \leq \epsilon$. By lemma (3.5.2), There exists a continuous linear extension h of $g|_{\ker f}$ to all X with $\|h\| \leq \epsilon$. Since $\ker f \subset \ker(g-h)$, there exists some scalar α such that $g - h = \alpha f$, see Theorem (3.5.3). Now note that

$$|\alpha| = \|\alpha f\| = \|g-h\| \leq \|g\| + \|h\| \leq 1+\epsilon \quad \because \|f\| = 1,\ \|g\| = 1, \|h\| \ \leq \epsilon$$

and that $0 < 1 - \varepsilon \leq \|g\| - \|h\| \leq \|g-h\| = |\alpha|$. So, $\alpha \neq 0$ and $|1-|\alpha|| \leq \epsilon$
Now if $\alpha > 0$ then $|1-\alpha| = |1-|\alpha||$ and thus

$$\|g-f\| = \|h + (\alpha-1)f\| \leq \|h\| + |1-\alpha|\|f\| \leq 2\epsilon$$

$$\therefore \quad g - h = \alpha f,\ g - f = \alpha f - f + h = (\alpha - 1)f + h$$

On the other hand $\alpha < 0$ then $1 + \alpha| = |1-|\alpha||$

$$\begin{aligned} \|g+f\| &= \|h + (1+\alpha)f\| \\ &\leq \|h\| + |1+\alpha| \qquad\qquad \because\ g-h = \alpha f,\ \ g+f = h + (1+\alpha)f,\ \|f\| = 1 \\ &\leq 2\epsilon \end{aligned}$$

□

Lemma 3.5.5. Let f and g be bounded linear functionals of norm one on a Banach space, and let $0 < \epsilon < \dfrac{1}{2}$ be given. If g is a positive linear functional with respect to the cone $K(f, \dfrac{\epsilon}{2+\epsilon})$, that is, if $g(x) \geq 0$ for all $x \in K(f, \dfrac{\epsilon}{2+\epsilon})$, then $\|f-g\| \leq 2\epsilon$.

Proof. Assume that f, g and $0 < \epsilon < \dfrac{1}{2}$ satisfying the stated property.
Choose some unit vector $x_0 \in X$ such that $f(x_0) > \dfrac{1+\epsilon}{2+\epsilon}$.
Note that for each $y \in \ker f$ satisfying $\|y\| \leq \dfrac{1}{\epsilon}$,

$$\|x_0 \pm y\| \leq \|x_0\| + \|y\| \leq 1 + \frac{1}{\epsilon} \leq \frac{2+\epsilon}{\epsilon} f(x_0) = \frac{2+\epsilon}{\epsilon} f(x_0 \pm y)$$

$$\because\ f(x_0) > \frac{1+\epsilon}{2+\epsilon} \qquad \text{i.e,}\ \ \frac{2+\epsilon}{\epsilon} f(x_0) > \frac{1+\epsilon}{\epsilon}$$

So, $f(x_0 \pm y) \geq \dfrac{\epsilon}{2+\epsilon}\|x_0 \pm y\|$. Therefore $x_0 \pm y \in \ker\left(f, \dfrac{\epsilon}{2+\epsilon}\right)$
implies $g(x_0 \pm y) \leq 0$, for all $x_0 + y \in \ker\left(f, \dfrac{\epsilon}{2+\epsilon}\right)$, $\quad \because\ g(x) \geq 0,\ x \in \ker\left(f, \dfrac{\epsilon}{2+\epsilon}\right)$
So, $|g(y)| \leq g(x_0) \leq 1$ for all $y \in \ker f$ with $\|y\| \leq \dfrac{1}{\epsilon}$. The latter easily yields $\|g|_{\mathbf{ker} f}\| \leq \epsilon$.
Now a glance at Lemma (3.5.5) guarantees that either $\|f-g\| \leq 2\epsilon$ is true.

To see this $\|f-g\| \leq 2\epsilon$ is true. Note that $2\epsilon < 1$ implies that there exists some unit vector $x \in X$ such that $f(x) > 2\epsilon = 2\epsilon\|x\| \geq \frac{\epsilon}{2+\epsilon}\|x\|$.
Thus $x \in K\left(f, \frac{\epsilon}{2+\epsilon}\right)$, so $g(x) \geq 0$. Consequently,

$$\|f+g\| \geq f(x) + g(x) \geq f(x) > 2\epsilon.$$

This prove that $\|f-g\| \leq 2\epsilon$ must be the case. □

Lemma 3.5.6. Let f be a norm one linear functional on a Banach space X, and let $0 < \partial < 1$ be given. If D is a nonempty closed bounded subset of X, then for each $d \in D$ there exists some $m \in D$ satisfying $D \cap [m + K(f,\partial)] = m$ and $m - d \in K(f,\partial)$.

Proof. Define a partial order on D by $x \geq y$ if $x - y \in K(f,\delta)$.
Now fix $d \in D$ and consider the non-empty set

$$D_d = \{x \in D : x \geq d\} = \{x \in D : x - d \in K(f,\delta)\}$$

Notice that a vector $m \in D$ satisfies

$$D \cap [m + k(f,\delta)] = \{m\}$$

and $m - d \in K(f,\delta)$ if and only if m is maximal element in D_d with respect to $\geq$. So, to complete the proof we must show that the partially ordered set $(D_d, \geq)$ has a maximal element.
By Zorn's lemma, it suffices to prove that every chain in D_d has an upper bound in D_d.
To this end, let $\mathcal{C}$ be chain in D_d.
So assume that for each $u \in \mathcal{C}$ there exists some $v \in \mathcal{C}$ with $v > u$. If we let $A = \mathcal{C}$ and $x_\alpha = \alpha$ for each $\alpha \in \mathcal{C}$, we can identify $\mathcal{C}$ with the increasing bounded net of real numbers, and hence a C Cauchy net. Since for any α and β we have either $x_\alpha \geq x_\beta$ or $x - \beta \geq \xi_\alpha$, it follows that $\delta\|x_\alpha - x_\beta\| \leq |f(x_\alpha) - f(x_\beta)|$ for all α and β. This implies that $\{\xi_\alpha\}$ is a Cauchy net. Since X is a Banach space this net converges in X, say to some $m \in X$. Clearly, $m \in D$ and since the cone $K(f,\delta)$ is closed, we get $m \geq x_\alpha$ for each α
That means m is an upper bound of the chain $\mathcal{C}$, and the proof is finished. □

We are now ready to state and prove the Bishop-Phelps Theorem.

Theorem 3.5.7 (Bishop-Phelps). For a closed convex subset K of a Banach space X we have the following:

(i) The set of support points of K is dense in the boundary of K.

(ii) If a addition K is norm bounded, then the set of bounded linear functionals on X that support K is dense in X'.

Proof. (i) Fix some $x_0 \in \partial C$ and let $\epsilon > 0$. Choose some $y_0 \in C$ such that

$$\|x_0 - y_0\| < \frac{\epsilon}{2}$$

by the separation corollary (3.1.11), there is non-zero continuous linear function f satisfying

$$f(y_0) > f(z) \;\; \forall z \in C.$$

without loss of generality we may normalize f so that its norm is one. that means

$$\|f\| = sup\{f(z) : \|z\| < 1\} = 1$$

Now let $K = K(f, \frac{1}{2})$ and $D = C \cap (x_0 + k)$. D is a nonempty closed convex set. If $x \in D$ then $x - x_0 \in K$. So,

$$\frac{1}{2}\|x - x_0\| \leq f(x - x_0) = f(x) - f(x_0) < f(y_0) - f(x_0) \leq \|y_0 - x_0\| < \frac{\epsilon}{2}$$

Hence, $\|x - x_0\| < \epsilon$ and thus $D \subset B_\epsilon(x_0)$. In particular, D is norm bounded set. According to Lemma (3.5.6), there exists some $m \in D$ such that $D \cap (m+K) = \{m\}$ and $m - x_0 \in K$. Clearly, $m \in C \cap (m + K)$. Now fix $x \in C \cup (m + K)$. Then there exists some $k \in K$ such that

$$x = m + k = x_0 + (m - x_0) + k \in x_0 + K.$$

and so $x \in C \cup (x_0 + K) = D$.
This implies that $x \in D \cup (m + K) = \{m\}$, that is $x = m$.
Consequently $C \cap (m + K) = \{m\}$ and so by Lemma (3.5.1), m is a support point of C. satisfying $\|x_0 - m\| < \epsilon$.

(ii) Now assume that C is a nonempty, closed, norm bounded, and convex subset of Banach space X. Fix $f \in X'$ with $\|f\| = 1$ and let $\epsilon > 0$. Pick some $o < \delta < \frac{1}{2}$ satisfying $2\delta < \epsilon$.
It suffices to show that there exists a norm one linear functional $g \in X'$ that supports C satisfying

$$\|f - g\| \leq 2\delta.$$

For simplicity, let $K = K(-f, \frac{\delta}{2+\delta})$.
By lemma (3.5.6) there exists some $m \in C$ such that $C \cap (m + K_0 = \phi$. So by the Separation Theorem(3.1.9), there exists some bounded linear functional $h \in X'$ of

norm one satisfying

$$h(c) \leq h(m+k)$$

for all $c \in C$ and all $k \in K$.

Now at a glance at Lemma (3.5.5) guarantees that $\|f-h\| \leq 2\delta < \epsilon$.

$\square$

Theorem 3.5.8 (Bishop-Phelps). Assume that A and B are two nonempty subsets of a Banach space X satisfying the following properties:

(a) A is closed and convex.

(b) B is bounded.

(c) There exists some $f \in X'$ with $\|f\| = 1$ such that $\sup f(A) < \inf f(B)$.

Then for each $\epsilon > 0$ we can find some $g \in X'$ with $\|g\| = 1$ and some $\alpha \in A$ so that $\|f-g\| \leq 2\epsilon$ and $g(a) = \sup g(A) < \inf g(B)$.

Proof. It is enough to consider $0 < \epsilon < \frac{1}{2}$. Let

$$\alpha = \sup f(A), \beta = \inf f(B),$$

and then fix γ such that $\alpha < \gamma < \beta$.

Now consider the nonempty bounded set

$$V = B + (\beta - \gamma)U$$

and note that $\inf f(V) = [\inf f(B)] - (\beta - \gamma) = \gamma$. Now let $\delta = \dfrac{2+\epsilon}{\epsilon}$ and then choose $u \in A$ such that

$$\alpha - f(u) < \frac{\gamma - \alpha}{2de}.$$

Now fix some $\theta > \max\left\{\dfrac{\gamma-\alpha}{2}, \sup_{y \in V}\|y-u\|\right\}$, put $k = \dfrac{2\delta\theta}{\gamma-\alpha}$ and note that $1 < \delta < k$, By Lemma (3.5.6) there exists some $\alpha_0 \in A$ such that $A \cup [a_0 + K(f, \frac{1}{2})] = \{a_0\}$ and $a_0 - u \in K(f, \frac{1}{2})$. We claim that

$$V \subset a_0 + K(f, \frac{1}{2})$$

To see this, note that for each $y \in V$ we have

$$\begin{aligned}\|y - a_0\| &\leq \|y - u\| + \|a_0 - u\| < \theta + \|a_0 - u\| \leq \theta + kf(a_0 - u)\\ &\leq \theta + k[\alpha - f(u)] < \theta + \frac{k(\gamma - \alpha)}{2\delta} = 2\theta\\ &< 2\delta\theta = k(\gamma - \alpha) \leq kf(y - a_0)\end{aligned}$$

Next, pick a nonzero linear functional $g \in X'$ with $\|g\| = 1$ such that

$$\begin{aligned}g(a_0) &= \sup g(A) \leq \inf g(a_0 + K(f, \frac{1}{k})) \leq \inf g(V)\\ &= [\inf g(B)] - (\beta - \gamma) < \inf g(B).\end{aligned}$$

Moreover, from $\inf g(a_0 + K(f, \frac{1}{k})) \leq g(a_0)$, it follows that the linear functional g is $K(f, \frac{1}{k})$-positive. Since $\frac{1}{k} < \frac{1}{\delta} = \frac{\epsilon}{2+\epsilon}$ implies $K(f, \frac{\epsilon}{2+\epsilon}) \subset K(f, \frac{1}{k})$, we see that g is also $\frac{\epsilon}{2+\epsilon}$-positive. But then a glance at Lemma (3.5.5) guarantees that $\|f - g\| \leq 2\epsilon$, and the proof is finished. □

The theorem has a number of interesting applications. The first one is a sharper Banach Space Version of the Strong Separtaion Hyperplane Theorem .

Corollary 3.5.9. Assume that A and B are two nonempty disjoint convex subsets of a Banach Space X is such that A is closed and B is weakly(and in particular norm) compact. Then there exist a non-zero linear funcitonal $g \in X'$ and vectors $a_0 \in A$ and $b_0 \in B$ such that

$$\sup g(A) = g(a_0) < g(b_0) = \inf g(B).$$

Proof. By Theorem (4) there is a nonzero functional $f \in X'$ satisfying $\sup f(A) < \inf f(B)$. Without loss of generality we may assume $\|f\| = 1$. The hypothesis of Theorem (4) are satisfying $g(a_0) = \sup g(A) < \inf g(B)$.Since B is weak* compact there is a point $b_0 \in B$ satisfying $g(b_0) = \inf g(B)$. □

Corollary 3.5.10. Every proper nonempty closed subset of a separable Banach space is the intersection of a countable collection of closed half spaces that support it.

Proof. Let C be a proper nonempty convex closed subset of a separable Banach sapace X and let $x - 1, x_2, ..$ be a countable subset of $X \backslash C$ that is norm dense in $X \backslash C$. For each n let $d_n = d(x_n, C) > 0$, the distance of x_n from C, and note that $C \cup (x_n + \frac{d_n}{2}U) = \phi$. Now, according to Theorem (4), for each n there exist some nonzero $g_n \in X'$ and some

$y_n \in C$ such that

$$g(y_n) = \sup g_n(C) < \inf g_n(x_n + \frac{d_n}{2}) \tag{3.7}$$

Next, take any $x \in X \backslash C$ and put $d = d(x, C) > 0$. Choose some x_k such that $\|x - x_k\| = \frac{d}{3}$. This implies that for $c \in C$ we have

$$\|c - x_k\| \geq \|c - x\| - \|x - x_k\| \geq d - \frac{d}{3} = \frac{2}{3}d.$$

Thus $d_k = \inf_{c \in C} \|c - x_k\| \geq \frac{2}{3}d$, which implies that $\|x - x_k\| < \frac{1}{2}d_k$. Consequently $x \in x_k + \frac{d_k}{2}U$ and from (4) we get $g_k(x) > \sup g_k(C)$ or

$$-g_k(x) < -g_k(y_k) = \inf[-g_k(C)],$$

and this leads to the desired conclusion. □

3.6 Extreme points

Definition 3.6.1. (i) An **extreme subset** of a (not necessarily convex) subset C of a vector space, is a nonempty subset F of C with the property that if x belongs to F it cannot be written as a convex combination of points of C outside F.
That is, if $x \in F$ and $x = \alpha y + (1 - \alpha)z$, where $0 < x < 1$ and $y.z \in C$, then $y, z \in F$.

(ii) A point x is an **extreme** point of C if the singleton $\{x\}$ is an extreme set. The set of extreme points of C is denoted $\mathcal{E}(C)$. That is, A vector x is an extreme point of C if it cannot be written as a strict convex combination of distinct points in C.

(iii) A **face** of a convex set C is a convex extreme subset of C.

Here is some examples.

- The extreme points of a closed disk are all the points on the circumference.
- The set of extreme points of a convex set is an extreme set - if it is nonempty.
- In $\mathbb{R}^n$, the extreme points of a convex polyhedron are its vertexes. All its faces and edges are extreme sets.
- The ray of pointed closed convex cone that are extreme sets are called **extreme rays**. For instance, the nonnegative axes are the extreme rays of the usual positive cone in $\mathbb{R}^n$.

The following useful property is easy to verify.

Lemma 3.6.2. A point a in a convex set C is an extreme point if and only if $C\backslash\{a\}$ is a convex set.

Remark 3.6.3. In general, the set of extreme points of a convex set K may be empty, and if nonempty, need not be closed.
For instance, the set C of all strictly positive functions on the unit interval is a convex subset of $\mathbb{R}^[0,1]$ without extreme points. To see this,
Let f be strictly positive. Then, $g = \frac{1}{2}f$ is also strictly positive and distinct from f, but

$$f = \frac{1}{2}g + \frac{1}{2}(f+g),$$

proving that f cannot be an extreme point of C.

As an example of a compact convex set for which the set of extreme points is not closed, consider the subset of $\mathbb{R}^3$.

Example 3.6.4. Consider the subset of $\mathbb{R}^3$

$$A = \{(x,y,0) \in \mathbb{R}^3 : x^2+y^2 \leq 1\} \cup \{(0,-1,1),(0,-1,-1)\}.$$

The convex hull of A is compact, but the set of extreme points of A is

$$\{(x,y,0) \in \mathbb{R}^3 : x^2+y^2 = 1\} \cup \{(0,-1,1),(0,-1,-1)\}\backslash\{(0,-1,0)\},$$

which is not closed. See figure (3.5)
You can verify the following properties of extreme subsets

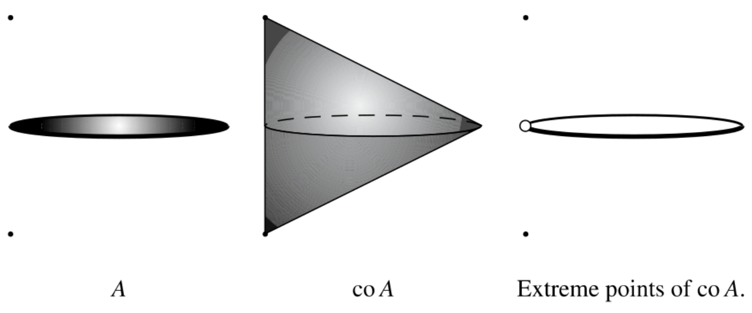


FIGURE 3.5

(i) An extreme subset of a set C is an extreme subset of C.

(ii) A nonempty intersection of a collection of extreme subset of the set C.

Lemma 3.6.5. The set of maximizers of a convex function is either set or empty. Likewise, the set of minimizers of a concave function is either on extreme set or empty.

Proof. Let $f : C \to \mathbb{R}$ be convex. Suppose f achieves a maximum on C. Put $M = \max\{f(x) : x \in C\}$ and let $F = \{x \in C : f(x) = M\}$. Suppose that $x = \alpha y + (1-\alpha)z \in F, 0 < \alpha < 1$, and $y, z \in C$. If $y \in F$, then $f(y) < M$,so

$$\begin{aligned} M = f(x) = f(\alpha y + (1-\alpha)z) &\leq \alpha f(y) + (1-\alpha)f(z) \\ &< \alpha M + (1-\alpha)M = M, \end{aligned}$$

a contradiction. Hence $y, z \in F$, so F is an extreme subset of C. □

Lemma 3.6.6. In a locally convex Hausdorff space, every compact extreme subset of a set C contains an extreme point of C.

Proof. Let C be a subset of some locally convex Hausdorff space and let F be a compact extreme subset of C. Consider the collection of sets

$$\mathcal{F} = \{G \subset F : G \text{ is a compact extreme subset of C}\}.$$

Since $F \in \mathcal{F}$, we have $\mathcal{F} \neq \phi$, and $\mathcal{F}$ is partially ordered by set inclusion.
The compactness of F (as expressed in terms of the infinite intersection property) guarantees that every chain in $\mathcal{F}$ has a nonempty intersection. Clearly, the intersection of extreme subsets of C is an extreme subset of C is it is nonempty.
Thus by Zorn's Lemma applies, and yields a minimal compact extreme subset of C included in F, call it G.
We claim that G is singleton.
To see this assume by way of contradiction that there exist $a, b \in G$ with $a \neq b$. By the Separation Corollary (4) there is a continuous linear functional f on X such that $f(a) > f(b)$.
Let M be the maximum value of f on G. Arguing as in the proof of Lemma (4), we see that the compact set

$$G_0 = \{c \in G : f(c) = M\}$$

is an extreme subset of G (and hence of C) and $b \in G_0$, contrary to the minimality of G. Hence G must be a singleton. Its unique element is an extreme point of C lying in F. □

3.7 Quasiconvexity

There are generalizations of convexity for functions that are commonly applied in economic theory and operations research.

Definition 3.7.1. A real function $f : K \to \mathbb{R}$ on a convex subset K of a vector space X is

- **Quasiconvex** if for all $x, y \in K$ and $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}$$

- **Strictly Quasiconvex**, If $f(\lambda + (1 - \lambda)y) < \max\{f(x), f(y)\}$ with $x \neq y$ and $\lambda \in (0, 1)$.
- **Quasiconcave**, if $-f$ is a quasiconvex function. Explicitly, $f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\}$ for all $x, y \in K$ and $\lambda \in [0, 1]$.
- **Strictly Quasiconcave**, if $-f$ is strictly quasiconvex.

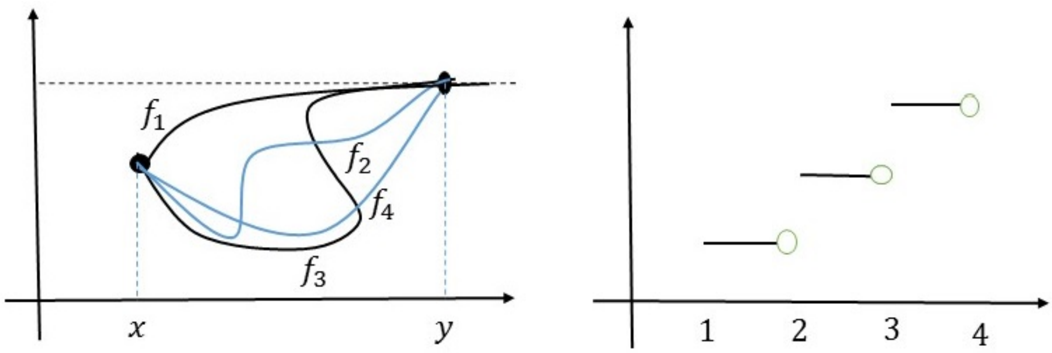


FIGURE 3.6: Graph of Quasiconvex function

Example 3.7.2.

Lemma 3.7.3. Every convex function is quasiconvex (and every concave function is quasiconcave)

Proof. Let $f : K \to \mathbb{R}$ be a convex function on convex subset K of a vector space X.
So, $f(x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad \forall x, y \in K$ and $\lambda \in [0, 1]$
Since, $\lambda f(x)+(1-\lambda)f(y) \leq \lambda \max\{f(x), f(y)\}+(1-\lambda) \max\{f(x), f(y)\} = \max\{f(x), f(y)\}$
Thus, $f(x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \leq \max\{f(x), f(y)\}$ and hence proved.
Similarly, we can show for concave function. $\square$

Lemma 3.7.4. For a real function $f : K \to \mathbb{R}$ on a convex set, the following statements are equivalent:

(1) The function f is quasiconvex.

(2) For each $\alpha \in \mathbb{R}$, the strict lower counter set $\{x \in K : f(x) < \alpha\}$ is a convex set.

(3) For each $\alpha \in \mathbb{R}$, the lower counter set $\{x \in K : f(x) \geq \alpha\}$ is a convex set.

Definition 3.7.5 (Explicitly Quasiconvex)**.** A real valued function g is *explicitly quasiconvex*, if it is a quasiconvex and in addition, $g(x) < g(y)$ implies $g(\lambda x + (1-\lambda)y) < g(y)$ for $0 < \lambda < 1$.

Theorem 3.7.6. A real-valued lower semicontinuous function on a compact space attains a maximum value, and the nonempty set of minimizers is compact. Similarly, an upper semicontinuous function on a compact set attains a maximum value, and the nonempty set of maximizers is compact.
For proof of this theorem see [4] (Theorem 2.43).

Corollary 3.7.7. Let K be nonempty compact convex subset of a locally convex Hausdorff space. Every upper semicontinuous explicitly quasiconvex function has a maximizer on K that is an extreme point of K

Proof. Let $f : C \to \mathbb{R}$ be an upper semicontinuous explicitly quasiconvex function. By theorem (3.7.6) the set F of maximizers of f is nonempty and compact.
Put $M = \max\{f(x) : x \in C\}$, So, $F = \{x \in C : f(x) = M\}$.
We wish to show that F is an extreme subset of C, that is, if x belongs to F, and $x = \alpha y + (1-\alpha)z$, where $0 < \alpha < 1$ and $y, z \in C$, then both y and z belongs to F. If say $y \notin F$, then $f(y) < M = f(x)$, so by quasiconvexity we have

$$M = f(x) \leq \max\{f(y), f(z)\},$$

which implies $f(x) = f(z) = M > f(y)$. other hand, since f is explicitly quasiconvex, and $f(y) < f(z)$, we must also have $f(x) < f(z)$, a contradiction. Therefore F is an extreme set.

By Lemma (4), F contains an extreme point of C. □

Bibliography

[1] H.G Eggleston, *Convexity*, Cambridge University,Press 1958.

[2] S.R Lay, *Convex sets and their applications, Wiley*, New Yark 1958.

[3] F.A Valentine, *Convex Sets*, Mcgraw Hill,1964.

[4] Charalanmbos D.Aliprantis Kim C.Border, *Infinite Dimensional Analysis*, Springer-Verlag Berlin Heidelberg 1999.

[5] J.Stoer and C.Witzgall, *Convexity and Optimization on Finite Dimensions*, Springer-Verlag, Heidelberg, New York 1970.

[6] Steven G Krantz,*Convex Analysis*,CRC Press,2014.

[7] W.H Flemming, *Functions of Several Variables*, -Wesley 1965,(elementary Theory).

[8] R.Phelps, *Convex Functions,Monotone Operators and Differentiability*, Springer-Verlag.

[9] R.T. Rockefeller, *Convex Analysis*, Princeton University Press,1970.

[10] L A suinow and A.J.Ellis, *Convexity Theory and its applications in Functional Analysis*, Academic Press, 1980.